Optimal quantum inference: using nonlinear convex analysis on noncommutative Banach spaces

Ryszard Paweł Kostecki

National Quantum Information Centre
Faculty of Mathematics, Physics, and Informatics, University of Gdańsk

Department of Mathematical Informatics, Nagoya University
13.03.2020

Plan

1. Noncommutative integration theory
2. Entropic projections as nonlinear state transformations
3. Brègman family of quantum relative entropies
4. Categories of Brègman nonexpansive operations
5. Brègman nonexpansive resource theories

前は闇, 後ろは輝く星座というのが作用素環の世界です.
竹崎正道, 作用素環への入り口

It is in the world of operator algebras that the forefront is dark and behind is a twinkling constellation.

Masamichi Takesaki, 2003, *Entrance to operator algebras*.

source of the photo: www.fuw.edu.pl/~kostecki/photos/_cqg

Probability theory:

- Underlying structure: measure space $(\mathcal{X}, \mu)$
- Main spaces: **Probabilistic models:** normalised subsets of:

$$\mathcal{M}(\mathcal{X}, \mu) \subseteq L_1(\mathcal{X}, \mu)^+ := \{p : \mathcal{X} \to \mathbb{R} \mid \int_{\mathcal{X}} \mu |p| < \infty, p \geq 0\}$$

- e.g. Gaussian models: $\{p(\chi, (m, s)) = \frac{1}{\sqrt{2\pi}s} \mathrm{e}^{-\frac{(\chi - m)^2}{2s^2}} \mid (m, s) \in \Theta \subseteq \mathbb{R} \times \mathbb{R}^+\}$.
- Observables: functions $f : \mathcal{X} \to \mathbb{R}$
- The mapping $L_1(\mathcal{X}, \mu) \times L_\infty(\mathcal{X}, \mu) \ni (p, f) \mapsto \int_{\mathcal{X}} \mu p f \in \mathbb{R}$ determines Banach space duality $L_1(\mathcal{X}, \mu)^\star \cong L_\infty(\mathcal{X}, \mu)$.
- convergence of integration: $\int_{\mathcal{X}} \mu p \sup_i(f_i) = \sup_i(\int_{\mathcal{X}} \mu p f_i)$

Quantum mechanics:

- Underlying structure: Hilbert space $\mathcal{H}$
- Main spaces: **Spaces of density matrices:** normalised subsets of:

$$\mathcal{M}(\mathcal{H}) \subseteq \mathcal{T}(\mathcal{H})^+ := \{\rho \in \mathfrak{B}(\mathcal{H}) \mid \mathrm{tr}_{\mathcal{H}}(|\rho|) < \infty, \rho \geq 0\}$$

- e.g. Gibbs states: $\{\mathrm{e}^{-\beta H} \mid \beta \in]0, \infty[\}$, for a fixed self-adjoint H.
- Observables: self-adjoint operators $x : \mathcal{H} \to \mathcal{H}$
- The mapping $\mathcal{T}(\mathcal{H}) \times \mathfrak{B}(\mathcal{H}) \ni (\rho, x) \mapsto \mathrm{tr}_{\mathcal{H}}(\rho x) \in \mathbb{C}$ determines Banach space duality $\mathcal{T}(\mathcal{H})^\star \cong \mathfrak{B}(\mathcal{H})$.
- convergence of integration: $\mathrm{tr}_{\mathcal{H}}(\rho \sup_i x_i) = \sup_i \mathrm{tr}_{\mathcal{H}}(\rho x_i)$.

Is there a joint generalisation of the above two settings?

W*-algebras and integration

Yes.

- A W*-algebra $\mathcal{N}$ [von Neumann'29, Sakai'56]:
  - a (noncommutative) algebra over $\mathbb{C}$ with unit $\mathbb{I}$,
  - with * operation s.t. $(xy)^* = y^*x^*$, $(x+y)^* = x^* + y^*$, $(x^*)^* = x$, $(\lambda x)^* = \lambda^* x^*$,
  - that is also a Banach space,
  - with $\cdot$, $+$, * continuous in the norm topology (implied by the condition $\|x^*x\| = \|x\|^2$),
  - such that there exists a Banach space $\mathcal{N}_\star$ satisfying the Banach space duality: $(\mathcal{N}_\star)^\star \cong \mathcal{N}$,
- Special cases:
  - if $\mathcal{N}$ is commutative
    then $\exists$ a measure space $(\mathcal{X}, \mu)$ s.t. $\mathcal{N} \cong L_\infty(\mathcal{X}, \mu)$ and $\mathcal{N}_\star \cong L_1(\mathcal{X}, \mu)$
  - if $\mathcal{N}$ is "type I"
    then $\exists$ a Hilbert space $\mathcal{H}$ s.t. $\mathcal{N} \cong \mathfrak{B}(\mathcal{H})$ and $\mathcal{N}_\star \cong \mathcal{T}(\mathcal{H})$.
- Hence, the element $\phi \in (\mathcal{N}_\star)^+$ provides a joint generalisation of probability density and of density operator.
- By means of embedding of $\mathcal{N}_\star$ into $\mathcal{N}^\star$, it is also an integral on $\mathcal{N}$.
- Hence, the subsets of $\mathcal{N}_\star^+$ can be considered as generic quantum state spaces.

Noncommutative integration on W*-algebras

- state is a function $\omega : \mathcal{N} \to \mathbb{C}$ s.t. $\omega \in \mathcal{N}^+_\star$
- faithful normal semifinite weight is a function $\omega : \mathcal{N}^+ \to [0, +\infty]$ s.t.
  $\omega(0) = 0$, $\omega(x + y) = \omega(x) + \omega(y)$, $\lambda > 0 \Rightarrow \omega(\lambda x) = \lambda\omega(x)$,
  $\forall x \in \mathcal{N}^+ \exists \mathcal{N}^+ \setminus \{0\} \ni y \leq x \quad \omega(y) < +\infty$
  $\omega(\sup \mathcal{F}) = \sup_{x\in\mathcal{F}} \omega(x)$ $\forall$ directed filters $\mathcal{F} \subseteq \mathcal{N}^+$,
  $\omega(x^* x) = 0 \Rightarrow x = 0 \forall x \in \mathcal{N}$
- trace is a weight s.t. $\omega(u^* x u) = \omega(x)$ $\forall$ unitary $u \in \mathcal{N}$
- every (faithful) state is a finite (faithful) normal weight.
- von Neumann–Murray['36-'43] classif. of W*-algebras: every $\mathcal{N}$ is a direct product of:
  - type I_n: isomorphic with $\mathfrak{B}(\mathcal{H})$, $\dim \mathcal{H} = n \in \mathbb{N} \cup \{+\infty\}$, so $(\mathfrak{B}(\mathcal{H}), \mathrm{tr}_\mathcal{H})$ by default
  - type II: not of type I, yet admitting f.n.s. trace τ (II_1 if finite, II_∞ otherwise)
  - type III: neither of the above (always admits f.n.s. weight ψ but not f.n.s. trace)

- Commutative integration:

	spatial representation (in general case)	algebraic formulation (general)
underlying object	localisable measure space: $(\mathcal{X}, \mho(\mathcal{X}), \mu)$	localisable boolean algebra: $\mathcal{A}$
L_p-spaces	$L_p(\mathcal{X}, \mho(\mathcal{X}), \mu)$	$Lp(\mathcal{A})$
states	$q \in \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \mu) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \mu)^+$	$\phi \in \mathcal{M}(\mathcal{A}) \subseteq L_1(\mathcal{A})^+$
expectations of observables	$L_\infty(\mathcal{X}, \mho(\mathcal{X}), \mu) \ni f \mapsto \int_\mathcal{X} \mu q f \in \mathbb{R}$	$\phi \in L_\infty(\mathcal{A}) \ni f \mapsto \phi(f) \in \mathbb{R}$

- Noncommutative integration:

	spatial representation (in type I case)	algebraic formulation (general)
underlying object	Hilbert sp. with std. trace: $(\mathcal{H}, \mathfrak{B}(\mathcal{H}), \mathrm{tr}_\mathcal{H})$	W*-algebra: $\mathcal{N}$
L_p-spaces	$\mathfrak{G}_p(\mathcal{H}) = L_p(\mathfrak{B}(\mathcal{H}), \mathrm{tr}_\mathcal{H})$	$L_p(\mathcal{N})$
states	$\rho \in \mathcal{M}(\mathcal{H}) \subseteq \mathfrak{G}_1(\mathfrak{B}(\mathcal{H}), \mathrm{tr}_\mathcal{H})^+ \cong \mathfrak{B}(\mathcal{H})^+_\star$	$\phi \in \mathcal{M}(\mathcal{N}) \subseteq L_1(\mathcal{N})^+ \cong \mathcal{N}^+_\star$
expectations of observables	$\mathfrak{B}(\mathcal{H}) = \mathfrak{G}_\infty(\mathfrak{B}(\mathcal{H}), \mathrm{tr}_\mathcal{H}) \ni x \mapsto \mathrm{tr}_\mathcal{H}(\rho x) \in \mathbb{C}$	$\mathcal{N} = L_\infty(\mathcal{N}) \ni x \mapsto \phi(x) \in \mathbb{C}$

GNS representation: from C^*-algebras to Hilbert spaces

- A representation of a C^*-algebra $\mathcal{A}$ is a pair $(\mathcal{H}, \pi)$, where $\pi : \mathcal{N} \to \mathfrak{B}(\mathcal{H})$ is a $*$-homomorphism: $\pi(\lambda_1 x + \lambda_2 y) = \lambda_1 \pi(x) + \lambda_2 \pi(y)$, $\pi(xy) = \pi(x)\pi(y)$, $\pi(x^*) = \pi(x)^*$.
- Gel'fand–Naĭmark'43–Segal'47 theorem: Every pair of C^*-algebra $\mathcal{A}$ and $\phi \in \mathcal{A}^{\star +}$ determines a unique representation $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$ where $\Omega_\omega \in \mathcal{H}_\omega$ is cyclic: $\mathcal{H}_\omega = \overline{\{\pi_\omega(x)\Omega_\omega \mid x \in \mathcal{A}\}}^{\|\cdot\|_{\mathcal{H}_\omega}}$. Proof goes by explicit construction:
  - $\langle x, y \rangle_\omega := \omega(x^* y) \; \forall x, y \in \mathcal{A}$
  - $\ker \omega = \{x \in \mathcal{A} \mid \omega(x^* x) = 0\}$
  - $\mathcal{H}_\omega := \overline{\mathcal{A}/\ker \omega}^{\langle\cdot,\cdot\rangle_\omega}$
  - $[\cdot]_\omega : \mathcal{A} \ni x \mapsto [x]_\omega \in \mathcal{A}/\ker \omega$
  - $\pi_\omega(x) : [y]_\omega \mapsto [xy]_\omega$
  - $\Omega_\omega := [\mathbb{I}]_\omega$
  - $\omega(x) = \langle \Omega_\omega, \pi_\omega(x)\Omega_\omega \rangle_\omega \; \forall x \in \mathcal{A}$
- Every representation of $\mathcal{A}$ can be decomposed as a countable or noncountable direct product of representations that are unitarily equivalent to GNS representation.
- If ω is faithful, then π_ω is a $*$-isomorphism, and Ω_ω is separating: $\pi_\omega(x)\Omega_\omega = 0 \Rightarrow \pi_\omega(x) = 0 \; \forall x \in \mathcal{A}$.
- If ψ is a n.s. weight on W^*-algebra $\mathcal{N}$, then the construction of $(\mathcal{H}_\psi, \pi_\psi)$ is the same, just with $\mathcal{A}$ replaced by the ideal $\mathfrak{n}_\psi := \{x \in \mathcal{N} \mid \psi(x^* x) < \infty\}$. There is no corresponding construction of a cyclic vector Ω_ψ. If ψ is also faithful, then π_ψ is $*$-isomorphism.
- If $\mathcal{N} = \mathfrak{B}(\mathcal{K})$ and $\psi = \mathrm{tr}_\mathcal{K}$ then $\mathcal{H}_\psi \cong \mathfrak{G}_2(\mathcal{K}, \mathrm{tr}_\mathcal{K})$, with $\langle x, y \rangle_\psi = \mathrm{tr}_\mathcal{K}(x^* y)$, and $\pi_\psi(x) = \mathfrak{L}_x$ (left multiplication).

Noncommutative $L_p(\mathcal{N},\tau)$ spaces

- GNS representation of a W*-algebra is called a von Neumann algebra [v.N.'29].
- A closed densely defined linear operator $x : \mathrm{dom}(x) \to \mathcal{H}$ with polar decomposition $x = v|x|$ is called to be affiliated with a von Neumann algebra $\mathcal{N}$ (acting on $\mathcal{H}$) iff $v \in \mathcal{N}$ and all spectral projections of $|x|$ belong to $\mathcal{N}$. [v.N.–Murray'36]
- $x : \mathrm{dom}(x) \to \mathcal{H}$ (as above) is called τ-measurable [Nelson'74] iff $\exists\lambda > 0$ $\tau(\pi_\tau^{-1}(P^{|x|}([\lambda,\infty[))) < \infty$ for a f.n.s. trace τ on a W*-algebra $\mathcal{N}$.
- [Nelson'74, Yeadon'75]: $\mathscr{M}(\mathcal{N},\tau) :=$ the space of all τ-measurable operators on $\mathcal{H}_\tau$ affiliated with $\pi_\tau(\mathcal{N})$ for any W*-algebra $\mathcal{N}$ with f.n.s. trace τ.
- [Segal'53] has introduced a more broad definition of measurability of operators. τ-measurable operators are always Segal-measurable.
- τ can be extended from $\mathcal{N}^+$ to $\mathscr{M}(\mathcal{N},\tau)^+$ by

$$\widetilde{\tau} : \mathrm{aff}(\pi_\tau(\mathcal{N}))^+ \ni x \mapsto \widetilde{\tau}(x) := \sup_{n\in\mathbb{N}} \left\{ \tau \circ \pi_\tau^{-1} \left(\int_0^n P^x(\lambda)\lambda \right) \right\} \in [0,\infty].$$

- [Segal'53, Ogasawara–Yoshinaga'55, Kunze'58, Nelson'74, Yeadon'75]: The map $\|\cdot\|_p : \mathscr{M}(\mathcal{N},\tau) \ni x \mapsto \|x\|_p := (\widetilde{\tau}(|x|^p))^{1/p} \in [0,\infty]$, $p \in [1,\infty[$ is a norm determining the Banach spaces:

$$L_p(\mathcal{N},\tau) := \{x \in \mathscr{M}(\mathcal{N},\tau) \mid \|x\|_p < \infty\}.$$

- $L_p(\mathcal{N},\tau)$ provide the concrete operator-theoretic model of abstract noncommutative L_p spaces, defined by Dixmier'53 as topological completions of the spaces $\{y \in \{x \in \mathcal{N}^+ \mid \tau(x) < \infty\} \mid \|x\|_p < \infty\}$ in the norm $\|x\|_p := \tau(|x|^p))^{1/p}$.

Noncommutative $L_p(\mathcal{N},\tau)$ spaces (II)

- For all $\gamma \in]0,1]$: $(x,y) \in L_{1/\gamma}(\mathcal{N},\tau) \times L_{1/(1-\gamma)}(\mathcal{N},\tau) \;\Rightarrow\; xy \in L_1(\mathcal{N},\tau)$,
- The duality $L_{1/\gamma}(\mathcal{N},\tau) \times L_{1/(1-\gamma)}(\mathcal{N},\tau) \ni (x,y) \mapsto [\![x,y]\!] := \tau(xy) \in \mathbb{R}$ determines an isometric isomorphism of Banach spaces $L_{1/\gamma}(\mathcal{N},\tau)^\star \cong L_{1/(1-\gamma)}(\mathcal{N},\tau)$.
- The noncommutative analogue of the Rogers–Hölder inequality reads $\|xy\|_1 \leq \|x\|_{1/\gamma}\|y\|_{1/(1-\gamma)} \;\; \forall (x,y) \in L_{1/\gamma}(\mathcal{N},\tau) \times L_{1/(1-\gamma)}(\mathcal{N},\tau)$.
- The space of finite rank operators over a Hilbert space,

$$\mathfrak{G}_{\mathrm{fin}}(\mathcal{H}) := \{x \in \mathfrak{B}(\mathcal{H}) \mid \dim \mathrm{ran}(x) < \infty\},$$

  allows to define: the space of Riesz'1917–Schauder'30 (= compact) operators over a Hilbert space $\mathcal{H}$,

$$\mathfrak{G}_0(\mathcal{H}) := \overline{\mathfrak{G}_{\mathrm{fin}}(\mathcal{H})}^{\|\cdot\|_{\mathfrak{B}(\mathcal{H})}};$$

- For any $p \in [1,\infty[$, the spaces $\mathfrak{G}_p(\mathcal{H})$ of (von Neumann'37–Schatten'50)'46'47 p-class operators over a Hilbert space $\mathcal{H}$ are defined as

$$\mathfrak{G}_p(\mathcal{H}) := \{x \in \mathfrak{G}_0(\mathcal{H}) \mid \|x\|_p := \mathrm{tr}_{\mathcal{H}}((x^*x)^{p/2})^{1/p} < \infty\},$$

  and they are Banach spaces with respect to the norm $\|\cdot\|_p$ for $p \in [1,\infty[$. In addition, one sets $\mathfrak{G}_\infty(\mathcal{H}) := \mathfrak{B}(\mathcal{H})$ with $\|x\|_\infty := \|x\|_{\mathfrak{B}(\mathcal{H})}$.
- $\mathfrak{G}_p(\mathcal{H}) = L_p(\mathfrak{B}(\mathcal{H}), \mathrm{tr}_{\mathcal{H}})$.

Commutative/measure-theoretic Orlicz spaces

- A function $\Upsilon : \mathbb{R} \to [0, \infty]$ is called Young['1912'26] iff $\Upsilon(0) = 0$, $\Upsilon(-x) = \Upsilon(x)$, $0 \not\equiv \Upsilon \not\equiv \infty$ on $]0, \infty[$, Υ is convex on $] - b_\Upsilon, b_\Upsilon[$ and $\lim_{x \to -b_\Upsilon} \Upsilon(x) = \Upsilon(b_\Upsilon)$, where $b_\Upsilon := \sup\{t > 0 \mid \Upsilon(t) < \infty\}$.
- Υ is (sometimes) called Orlicz iff it is Young, continuous and nondecreasing on $\mathbb{R}^+$.
- Every Orlicz function Υ defines a Banach space [Orlicz'32'36]

$$L_\Upsilon(\mathcal{X}, \mho(\mathcal{X}), \mu; \mathbb{R}) = \{f \in L_0(\mathcal{X}, \mho(\mathcal{X}), \mu; \mathbb{R}) \mid \exists \lambda > 0 \ \int_{\mathcal{X}} \mu \Upsilon(\lambda |f|) < \infty\}$$

with the norm [Morse–Transue'50, Nakano'51, Luxemburg'55]

$$\|x\|_\Upsilon := \inf\{\lambda > 0 \mid \int_{\mathcal{X}} \mu \Upsilon(\lambda^{-1} x) \leq 1\} \in \mathbb{R}^+$$

- Zaanen'49–Luxemburg'55: Extension to any Young function.
- The spaces $L_p(\tilde{\mu})$ for $p \in [1, \infty[$ can be defined as Orlicz spaces $L_\Upsilon(\tilde{\mu})$ with $\Upsilon(x)$ given by the Orlicz functions: $\frac{|x|^p}{p}$ or $|x|^p$.
- The space $L_\infty(\mu)$ can be determined as an Orlicz space $L_{\Upsilon_\infty}(\mu)$, where $\Upsilon_\infty : \mathbb{R} \to [0, \infty]$, defined by Young function which is not an Orlicz function:

$$\Upsilon_\infty(x) := \begin{cases} 0 & : \ x \in [0, 1[\\ +\infty & : \ x > 1 \\ \Upsilon_\infty(-x) & : \ x < 0. \end{cases}$$

Noncommutative Orlicz spaces

- Rao'71: Υ is an Orlicz function, $\Upsilon(|x|)$ for $x \in \mathfrak{B}(\mathcal{H})$ is understood in terms of the spectral representation, the n.c. Orlicz space is defined as
$$\mathfrak{G}_\Upsilon(\mathcal{H}) := \left\{x \in \mathfrak{B}(\mathcal{H}) \mid \|x\|_\Upsilon := \inf\left\{\lambda > 0 \mid \mathrm{tr}_\mathcal{H}\left(\Upsilon\left(\lambda^{-1}|x|\right)\right) \leq 1\right\} < \infty\right\}.$$
- Hardy–Littlewood'30 (resp., Grothendieck'55–Sonis'71/Ovchinnikov'70–Yeadon'75): The rearrangement of $f \in L_0(\mathcal{X},\mu)^+$ (resp., $x \in \mathscr{M}(\mathcal{N},\tau)^+$) is defined as:
$$\mathbf{R}^\mu_f : [0,\infty[\ni t \mapsto \inf\{s \geq 0 \mid \ \mu\{\lambda \mid f(\lambda) > s\} \leq t\} \in [0,\infty],$$
$$\mathbf{R}^\tau_x : [0,\infty[\ni t \mapsto \mathbf{R}^\tau_x(t) := \inf\{s \geq 0 \mid \tau(P^x(]s,+\infty[) \leq t\} \in [0,\infty].$$
- Muratov'78'79: Construction of Orlicz spaces $L_\Upsilon(\mathcal{N},\tau)$ of Segal-measurable operators affiliated with (type I and) II_1 W*-algebras $\mathcal{N}$ (i.e., with f.n. finite traces τ), using $\mathbf{R}^\tau_x$, with Orlicz Υ satisfying $\lim_{t\to\infty}\frac{\Upsilon(t)}{t} = \infty$, $\lim_{t\to 0}\frac{\Upsilon(t)}{t} = 0$.
- Muratov'79, Kosaki'81, Ciach'83, Fack–Kosaki'86: $\mathbf{R}^\tau_{f(x)}(t) = f(\mathbf{R}^\tau_x(t))\ \ \forall t \in \mathbb{R}^+$ for any continuous increasing f on $[0,\infty[$ with $f(0) \geq 0$.
- Fack–Kosaki'86 ($\tilde{\tau}$ denotes an extension of τ from $\mathcal{N}$ to $\mathscr{M}(\mathcal{N},\tau)$):
$$\tilde{\tau}(f(x)) = \int_0^\infty \mathrm{d}t\, f(\mathbf{R}^\tau_x(t)).$$
- Kunze'90: For any W*-algebra $\mathcal{N}$ with f.n.s. trace τ (i.e. $\mathcal{N}$ is type I or type II):
$$L_\Upsilon(\mathcal{N},\tau) = \{x \in \mathscr{M}(\mathcal{N},\tau) \mid \exists\lambda > 0\ \ \tilde{\tau}(\Upsilon(\lambda|x|)) < \infty\}$$
with $\|\cdot\|_\Upsilon : \mathscr{M}(\mathcal{N},\tau) \ni x \mapsto \inf\{\lambda > 0 \mid \tilde{\tau}(\Upsilon(\lambda^{-1}|x|)) \leq 1\}$ and Orlicz Υ.

Symmetric ($\equiv$ rearrangement-invariant) operator spaces

- Semënov'64: A Banach space $(L, \|\cdot\|_L)$ which is a linear subspace of $L_0(\mathcal{X}, \mu)$ is called a symmetric function space iff

$$(f \in L, g \in L_0(\mathcal{X}, \mu), \mathbf{R}_{|g|} = \mathbf{R}_{|f|}) \;\Rightarrow\; (g \in L, \|g\|_L = \|f\|_L).$$

- Ovchinnikov'70'71 & Yeadon'75: consider a topological *-subalgebra $\mathfrak{C}_0(\mathcal{N}, \tau)$ of $\mathscr{M}(\mathcal{N}, \tau)$, consisting of τ-compact operators:

$$\mathfrak{C}_0(\mathcal{N}, \tau) := \{x \in \mathscr{M}(\mathcal{N}, \tau) \mid \forall\lambda > 0 \;\; \tau(\pi_\tau^{-1}(P^{|x|}([\lambda, \infty[))) < \infty\}.$$

- Ovchinnikov'70'71: A symmetric operator space is defined as a Banach space $(L(\mathcal{N}, \tau), \|\cdot\|_{L(\mathcal{N},\tau)})$, which is a linear subspace of $\mathfrak{C}_0(\mathcal{N}, \tau)$ and satisfies

$$(f \in L(\mathcal{N}, \tau), g \in \mathfrak{C}_0(\mathcal{N}, \tau), \mathbf{R}^\tau_{|g|} = \mathbf{R}^\tau_{|f|}) \;\Rightarrow\; (g \in L(\mathcal{N}, \tau), \|g\|_{L(\mathcal{N},\tau)} = \|f\|_{L(\mathcal{N},\tau)}).$$

- This includes completely the theory of symmetrically normed ideals of compact operators in $\mathfrak{B}(\mathcal{H})$, as developed by (von Neumann'37–Schatten'46'50'60)'46'47, Macaev'61, Gokhberg–Kreĭn'61'65'67, Russu'69.
- Yeadon'80: a bit different definition of symmetric operator space, starting from interpolation spaces.
- Medzhitov'87: General case: replace $\mathfrak{C}_0(\mathcal{N}, \tau)$ [$\forall\lambda > 0$] by $\mathscr{M}(\mathcal{N}, \tau)$ [$\exists\lambda > 0$].
- <u>Theorem</u>: [Kalton–Sukochev'08] (earlier versions: [Yeadon'80, Dodds$^{\otimes 2}$–de Pagter'89, Sukochev–Chilin'90]): Given a symmetric function space $(L, \|\cdot\|_L)$,

$$L(\mathcal{N}, \tau) := \{x \in \mathscr{M}(\mathcal{N}, \tau) \mid \mathbf{R}^\tau_x \in L\}, \;\; \text{with a norm} \;\; x \mapsto \|\mathbf{R}^\tau_x\|_L$$

is a symmetric operator space.

N.c. Orlicz spaces as symmetric operator spaces

- In this sense, Dodds$^{\otimes 2}$–de Pagter'89 & Sukochev–Chilin'90 allow ("implicitly contain") a definition of noncommutative Orlicz spaces which is equivalent to Kunze'90: for any Orlicz Υ:

$$L_\Upsilon(\mathcal{N},\tau) := \{x \in \mathscr{M}(\mathcal{N},\tau) \mid \mathbf{R}^\tau_x \in L_\Upsilon(\mathbb{R}^+, \mho_{\mathrm{Borel}}(\mathbb{R}^+), \mathrm{d}\lambda)\}$$

- Arazy'81 (type I)/Sukochev'86 (type II$_1$)/Dodds$^{\otimes 2}$–de Pagter'93'14 (II$_\infty$, if L is strongly symmetric, i.e. $\int_0^t \mathrm{d}t\mathbf{R}^\mu_x(r) \leq \int_0^t \mathrm{d}r\mathbf{R}^\mu_y(r)\ \forall t \geq 0 \Rightarrow \|x\|_L \leq \|y\|_L\ \forall x, y \in L$): L is reflexive $\Rightarrow L(\mathcal{N},\tau)$ is reflexive.
- Krygin–Sukochev–Chilin'91: L is uniformly convex $\Rightarrow L(\mathcal{N},\tau)$ is uniformly convex.
- There are further theorems relating correspondingly other geometric properties of L and $L(\mathcal{N},\tau)$ (also in the opposite direction).
- Burkill'28: Υ is said to satisfy $\triangle_2$ condition iff $\exists\lambda > 0\ \ \forall x \geq 0\ \ \Upsilon(2x) \leq \lambda\Upsilon(x)$.
- Luxemburg'55: For nonatomic $(\mathcal{X},\mu)$ with $\mu(\mathcal{X}) = \infty$, $L_\Upsilon(\mathcal{X},\mu)$ is reflexive iff both Υ and Υ^{Y} satisfy $\triangle_2$ condition, where

$$\Upsilon^{\mathsf{Y}}(y) := \sup_{x\geq 0}\{x|y| - \Upsilon(x)\}.$$

- Kamińska'82: For nonatomic $(\mathcal{X},\mu)$ with $\mu(\mathcal{X}) = \infty$, $(L_\Upsilon(\mathcal{X},\mu), \|\cdot\|_\Upsilon)$ is uniformly convex iff both Υ and Υ^{Y} satisfy $\triangle_2$ condition and are uniformly convex.
- The characterisation of uniform convexity (and other geometric properties) of commutative (resp., noncommutative) Orlicz spaces differs dependently on:
  (1) whether $(\mathcal{X},\mu)$ is atomic (resp., type I), nonatomic finite (resp., type II$_1$), or nonatomic infinite (resp., type II$_\infty$);
  (2) the choice of norm (apart from $\|\cdot\|_\Upsilon$, there are also Orlicz, p-Amemiya, and other...).

KMS states: a generalisation of Gibbs states to any W*-algebra

- For a W*-algebra $\mathcal{N}$, a group homomorphism $\alpha : \mathbb{R} \ni t \mapsto \alpha_t \in \mathrm{Aut}(\mathcal{N})$ is called weakly-$\star$ continuous iff $t \mapsto \phi(\alpha_t(x))$ is continuous $\forall (x, \phi) \in \mathcal{N} \times \mathcal{N}_\star^+$.
- $\mathcal{N}^\alpha_\infty := \{x \in \mathcal{N} \mid \exists!$ extension of α to an analytic function $\mathbb{C} \ni z \mapsto \alpha_z(x) \in \mathbb{C}\}$ is a $*$-subalgebra of $\mathcal{N}$.
- Kubo'57–Martin–Schwinger'59, Haag–Hugenholtz–Winnink'67, Kastler–Pool–Poulsen'69: given $(\mathcal{N}, \alpha)$ as above, with $\mathcal{N}^\alpha_\infty \neq \varnothing$, the state $\omega \in \mathcal{N}_\star^+$ is said to be: KMS w.r.t. α at β iff $\beta \in \mathbb{R} \setminus \{0\}$ and

$$\omega(y\alpha_{z+\mathrm{i}\beta}(x)) = \omega(\alpha_z(x)y) \;\; \forall x \in \mathcal{N}^\alpha_\infty \;\; \forall y \in \mathcal{N} \;\; \forall z \in \mathbb{C},$$

or $\beta = 0$ and $\omega(xy) = \omega(yx)$ and

$$\omega(\alpha_t(x)) = \omega(x) \;\; \forall x \in \mathcal{N}. \tag{1}$$

- every KMS state satifies (**??**)
- The set of all KMS states for fixed (α, β) is convex and compact in the weak-$\star$ topology of $\mathcal{N}^\star$
- The KMS condition makes sense also for n.s. weights ψ on $\mathcal{N}$, just under constraint of the domain from $\mathcal{N}$ to $\mathfrak{n}_\psi \cap \mathfrak{n}_\psi^*$.
- If $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, $\dim \mathcal{H} < \infty$, $\alpha_t = \mathrm{e}^{\mathrm{i}th}(\cdot)\mathrm{e}^{-\mathrm{i}th}$, $t \in \mathbb{R}$, $h \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}}$, then $\omega \in \mathcal{N}_\star^+$ satisfies KMS for α and β **iff** $\omega = \mathrm{tr}_\mathcal{H}(\rho\,\cdot\,)$ with $\rho = \mathrm{e}^{-\beta h}$:

$$\mathrm{tr}_\mathcal{H}(\rho xy) = \mathrm{tr}_\mathcal{H}(\mathrm{e}^{-\beta h}x\mathrm{e}^{\beta h}\mathrm{e}^{-\beta h}y) = \mathrm{tr}(\mathrm{e}^{-\beta h}y\alpha_{\mathrm{i}\beta}(x)) = \mathrm{tr}(\rho y\alpha_{i\beta}(x)).$$

Tomita–Takesaki modular theory

- Tomita'67 theorem: Every W^*-algebra $\mathcal{N}$ and f.n.s. weight ψ uniquely determines:
  - a weakly-$\star$ continuous group homomorphism $\sigma^\psi : \mathbb{R} \to \mathrm{Aut}(\pi_\psi(\mathcal{N}))$,
  - an antilinear $*$-isomorphism $j_\psi : \pi_\psi(\mathcal{N}) \to [\pi_\psi(\mathcal{N})]^\bullet$,

  such that
  - $\sigma^\psi_t : x \mapsto \Delta^{it}_\psi x \Delta^{-it}_\psi$, where $\{\Delta^{it}_\psi \mid t \in \mathbb{R}\}$ is a strongly continous group of unitaries in $\mathfrak{B}(\mathcal{H}_\psi)$,
  - $j_\psi : x \mapsto J_\psi x J_\psi$, where $J^*_\psi = J^{-1}_\psi = J_\psi$ and $J^2_\psi = \mathbb{I}_{\mathcal{H}_\psi}$,
  - $J_\psi \Delta^{1/2}_\psi [x]_\psi = [x^*]_\psi$.
- The group $\sigma^\psi : t \mapsto \Delta^{it}_\psi(\cdot)\Delta^{-it}_\psi$ is called the group of modular automorphisms.
- $\Delta_\psi =: \mathrm{e}^{-K_\psi}$ is a positive unbounded linear operator on $\mathcal{H}_\psi$, and defines unbounded modular hamiltonian $K_\psi = K^*_\psi$.
- If ψ is a faithful normal state, then $\Delta^{it}_\psi \Omega_\psi = \Omega_\psi$ and $K_\psi \Omega_\psi = 0$.
- Winnink'70–Takesaki'70 theorem:
  - ω is a unique element of $\mathcal{N}^+_\star$ that satisfies the KMS condition for $\pi^{-1}_\omega \circ \sigma^\omega$ and $\beta = 1$,
  - $\pi^{-1}_\omega \circ \sigma^\omega$ is the unique strongly continuous 1-parameter group for which ω satisfies the KMS condition with $\beta = 1$.
- Under notational redefinition of $\pi^{-1}_\omega \circ \sigma^\omega \circ \pi_\omega$ as σ^ω, the KMS condition implies $\omega(x) = \omega(\sigma^\omega(x))$ $\forall x \in \mathcal{N}$ and

$$\mathcal{N}_{\sigma^\omega} := \{x \in \mathcal{N} \mid \sigma^\omega_t(x) = x\} = \{x \in \mathcal{N} \mid \omega(xy) = \omega(yx) \;\; \forall x, y \in \mathcal{N}\}.$$

- The W.-T. thm. holds also for f.n.s. weights. Also: ψ is f.n.s. trace iff $\Delta_\psi = \pi_\psi(\mathbb{I})$.
- Hence, modular theory characterises nontraciality of a weight, and plays a fundamental role in the structure of type III W^*-algebras.

Tomita–Takesaki modular theory: type I_n example

- Even for type I_n case the modular automorphism of nontracial states is nontrivial.
- If $\mathcal{N}_{\star 0} \neq \emptyset$ (=: $\mathcal{N}$ is countably finite), then modular theory can be equivalently stated in terms of a von Neumann algebra $\mathcal{A}$ on a Hilbert space $\mathcal{H}$, together with a vector $\Omega \in \mathcal{H}$ that is cyclic and separating for $\mathcal{A}$, where $(\mathcal{H}, \mathcal{A}, \Omega) = (\mathcal{H}_\omega, \pi_\omega(\mathcal{N}), \Omega_\omega)$ and $\omega \in \mathcal{N}_{\star 0}$.
- Let $\mathcal{N} = \mathfrak{B}(\mathbb{C}^n)$, $\psi = \mathrm{tr}$, and consider representation $\pi := \mathfrak{L}$ of $\mathcal{N}$ on $\mathcal{H}_\psi$.
- For any faithful $\omega(\cdot) = \mathrm{tr}(\rho\,\cdot\,) = \mathrm{tr}(\sqrt{\rho}\,\cdot\,\sqrt{\rho})$, the vector $\sqrt{\rho}$ is cyclic and separating for $\pi(\mathcal{N})$ on $\mathcal{H}_\psi$, so π is unitarily equivalent to the GNS representation π_ω.
- Assuming that $\alpha_t := \mathrm{e}^{\mathrm{i}ht}(\cdot)\mathrm{e}^{-\mathrm{i}ht} \in \mathrm{Aut}(\mathcal{N})$ implements the action of the modular automorphism of $\mathcal{N}$ w.r.t. ω, we will find the explicit form of $\Delta_\omega^{\mathrm{i}t}$.
- Representation of α_t by $u(t) := \mathrm{e}^{\mathrm{i}\pi(h)t}$ does not satisfy $u(t)\sqrt{\rho} = \sqrt{\rho}$.
- This condition holds for $\Delta_{\sqrt{\rho}}^{\mathrm{i}t} = \mathrm{e}^{\mathrm{i}(\pi(h) - J\pi(h)J)t} = \mathrm{e}^{\mathrm{i}(\mathfrak{L}_h - \mathfrak{R}_h)t}$, with $\mathfrak{R}_x : y \mapsto yx^*$.
- Hence, given a strictly positive ρ assumed to satisfy KMS condition w.r.t. $\mathrm{e}^{\mathrm{i}ht}(\cdot)\mathrm{e}^{-\mathrm{i}ht}$ at $\beta = 1$, its modular hamiltonian is $K_\omega = \mathfrak{L}_h - \mathfrak{R}_h$.
- In particular, for $\rho = \mathrm{e}^{-\beta h}$: $\Delta_\rho = \Delta_{\mathrm{tr}(\rho\,\cdot\,)} = \mathfrak{L}_\rho(\mathfrak{R}_\rho)^{-1} = \mathfrak{L}_\rho \mathfrak{R}_{\rho^{-1}}$.

Relative modular theory

- Tomita's theorem relies on the following fact: if $\mathcal{N}$ is a W*-algebra, and ω is a f.n.s. weight on $\mathcal{N}$, then $\exists!$ densely defined closeable antilinear operator
$$R_\omega : [\mathfrak{n}_\omega \cap \mathfrak{n}_\omega^*]_\omega \ni [x]_\omega \mapsto [x^*]_\omega \in \mathcal{H}_\omega = \mathfrak{n}_\omega.$$
- Its closure $\overline{R}_\omega$ has a unique polar decomposition $\overline{R}_\omega = J_\omega \Delta_\omega^{1/2} = \Delta_\omega^{-1/2} J_\omega$.
- Araki'73–Connes'73'74–Digernes'75 relative modular theory: For any W*-algebra $\mathcal{N}$, n.s. weight ϕ and f.n.s. weight ω (both on $\mathcal{N}$), the map
$$R_{\phi,\omega} : [x]_\omega \mapsto [x^*]_\phi \;\; \forall x \in \mathfrak{n}_\omega \cap \mathfrak{n}_\phi^*$$
is a densely defined, closable antilinear operator. Its closure has a unique polar decomposition $\overline{R}_{\phi,\omega} = J_{\phi,\omega}\Delta_{\phi,\omega}^{1/2}$. From definition: $\Delta_{\omega,\omega} = \Delta_\omega$, $J_{\omega,\omega} = J_\omega$.
- $\Delta_{\phi,\omega}$ is positive self-adjoint unbounded, with $\mathrm{supp}(\Delta_{\phi,\omega}) = \mathrm{supp}(\omega)\mathcal{H}_\phi$, called a relative modular operator. It can be seen as a general form of noncommutative Radon–Nikodým quotient of two weights. (The condition $\phi \ll \omega$ follows from faithfulness of ω.) Araki'73: relative modular hamiltonian $K_{\phi,\omega} := -\log \Delta_{\phi,\omega}$.
- The ultrastrongly-$\star$ continous 1-parameter family of partial isometries
$$\mathbb{R} \ni t \mapsto [\phi : \omega]_t := \Delta_{\phi,\omega}^{\mathrm{i}t}\Delta_\omega^{-\mathrm{i}t} \in \mathrm{supp}(\phi)\mathcal{N}$$
is called the Connes' cocycle.
- If $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, $\phi = \mathrm{tr}_\mathcal{H}(\rho\,\cdot\,)$, $\omega = \mathrm{tr}_\mathcal{H}(\sigma\,\cdot\,)$, then $\Delta_{\phi,\omega} = \mathfrak{L}_\rho \mathfrak{R}_{\sigma^{-1}}$, $K_{\phi,\omega} = -\log(\mathfrak{L}_\rho \mathfrak{R}_{\sigma^{-1}})$.

Natural cone [Woronowicz'72'74, Connes'72'74, Haagerup'73'75, Araki'74]

- Motivation: generalise $\rho \mapsto \sqrt{\rho}$ to arbitrary W*-algebra $\mathcal{N}$ and n.s.f. weight ω.

$$\mathcal{H}^\natural_\omega = \overline{\bigcup_{x\in\mathfrak{n}_\omega\cap\mathfrak{n}^*_\omega} \{\pi_\omega(x)J_\omega[x]_\omega\}}^{\mathcal{H}_\omega} \quad \text{satisfies:}$$

- $\mathcal{H}^\natural_\omega$ is closed, convex, and self-polar: $\mathcal{H}^\natural_\omega = \{\zeta\in\mathcal{H}_\omega \mid \langle\xi,\zeta\rangle_{\mathcal{H}_\omega}\geq 0\;\;\forall\xi\in\mathcal{H}^\natural_\omega\}$,
- hence: it is pointed ($\mathcal{H}^\natural_\omega\cap -\mathcal{H}^\natural_\omega=\{0\}$), $\mathrm{span}_{\mathbb{C}}\mathcal{H}^\natural_\omega=\mathcal{H}_\omega$, and determines a partial order on $\mathcal{H}^{\mathrm{sa}}_\omega := \{\xi\in\mathcal{H}_\omega \mid J_\omega\xi=\xi\}$ by: $\xi\leq\zeta \iff \xi-\zeta\in\mathcal{H}^\natural_\omega\;\forall\xi,\zeta\in\mathcal{H}^{\mathrm{sa}}_\omega$,
- $\forall\phi\in\mathcal{N}^+_\star\;\exists!\zeta_\omega(\phi)\in\mathcal{H}^\natural_\omega\;\forall x\in\mathcal{N}\;\;\phi(x)=\langle\zeta_\omega(\phi),\pi_\omega(x)\zeta_\omega(\phi)\rangle_{\mathcal{H}_\omega}$,
- the map $\mathcal{N}^+_\star\ni\phi\mapsto\zeta_\omega(\phi)\in\mathcal{H}^\natural_\omega$ is order preserving,
- the map $\zeta^\natural_\omega:\mathcal{H}^\natural_\omega\to\mathcal{N}^+_\star$, defined by the condition

$$\zeta^\natural_\omega(\xi)(x)=\langle\xi,\pi_\omega(x)\xi\rangle_{\mathcal{H}_\omega}\quad\forall x\in\mathcal{N}\;\;\forall\xi\in\mathcal{H}^\natural_\omega,$$

  is a bijective norm continuous homeomorphism with $(\zeta^\natural_\omega)^{-1}=\zeta_\omega$,
- $\omega\in\mathcal{N}^+_{\star 0}\neq\varnothing\Rightarrow\zeta_\omega(\omega)=\Omega_\omega$; if ω is also tracial, then $\mathcal{H}^\natural_\omega=\overline{\pi_\omega(\mathcal{N})^+\Omega_\omega}^{\mathcal{H}_\omega}$,
- if $\mathcal{N}=\mathfrak{B}(\mathcal{K})$ and $\omega=\mathrm{tr}_{\mathcal{K}}$, then $\pi_\omega=\mathfrak{L}$, $\mathcal{H}^\natural_\omega=\mathfrak{G}_2(\mathcal{K})^+$, $J_\omega(\xi)=\xi^*$, $\zeta_\omega:\mathfrak{G}_1(\mathcal{K})^+\ni\rho\mapsto\rho^{1/2}\in\mathfrak{G}_2(\mathcal{K})^+$.
- Haagerup'75: The notion of standard representation, axiomatically characterising the above properties independently of the choice of n.s.f. weight.
- Kosaki'80: canonical representation/cone, given by a positive cone $L_2(\mathcal{N})^+$ for an arbitary W*-algebra $\mathcal{N}$, so that $\zeta:\mathcal{N}^+_\star\cong L_1(\mathcal{N})^+\ni\phi\mapsto\phi^{1/2}\in L_2(\mathcal{N})^+$.

Kosaki'80: Noncommutative $L_p(\mathcal{N})$ spaces

The approach of Kosaki is based on the use of polar decomposition of elements of $\mathcal{N}_\star$ in terms of relative modular operator. For $\phi_1, \phi_2 \in \mathcal{N}_\star$ with polar decompositions $\phi_1 = |\phi_1|(\cdot\, u_1)$ and $\phi_2 = |\phi_2|(\cdot\, u_2)$, $p \in [1, \infty[$, and $\lambda = \mathrm{e}^{\mathrm{i}r}|\lambda| \in \mathbb{C}$ with $r \in [0, 2\pi[$, consider the addition, multiplication and $*$ operations on $\mathcal{N}_\star$ given by

1) $\phi_1^{1/p} + \phi_2^{1/p} := (\varphi(\cdot\, u))^{1/p}$, where $\varphi \in \mathcal{N}_\star^+$ and a partial isometry u with $\mathrm{supp}(\varphi) = u^* u$ are determined by
$$u\Delta^{1/p}_{\varphi,|\phi_1|+|\phi_2|} := u_1\Delta^{1/p}_{|\phi_1|,|\phi_1|+|\phi_2|} + u_2\Delta^{1/p}_{|\phi_2|,|\phi_1|+|\phi_2|},$$
2) $\lambda \cdot \phi_1^{1/p} := (|\lambda|^p|\phi_1|(\cdot\, \mathrm{e}^{\mathrm{i}r} u))^{1/p}$,
3) $(\phi_1^{1/p})^* := (\varphi(\cdot\, u))^{1/p}$, where $\varphi \in \mathcal{N}_\star^+$ and a partial isometry u with $\mathrm{supp}(\varphi) = u^* u$ are determined by
$$u\Delta^{1/p}_{\varphi,|\phi_1|} := (u_1\Delta^{1/p}_{|\phi_1|})^*.$$

$\phi^{1/p}$ is understood here as a *symbol* referring to the element ϕ of $\mathcal{N}_\star$ subject to the above operations. The set $\mathcal{N}_\star^+$ equipped with the above structure becomes a vector space with involution *, and will be denoted by $\mathscr{M}^p(\mathcal{N})$. The map
$$\|\cdot\|_p : \mathscr{M}^p(\mathcal{N}) \ni \phi^{1/p} \mapsto \left\|\phi^{1/p}\right\|_p := (|\phi|(\mathbb{I}))^{1/p} = \|\phi\|^{1/p}_{\mathcal{N}_\star}$$
defines a norm on $\mathscr{M}^p(\mathcal{N})$, with respect to which $\mathscr{M}^p(\mathcal{N})$ is Cauchy complete. The Banach spaces $(\mathscr{M}^p(\mathcal{N}), \|\cdot\|_p)$ are denoted by $L_p(\mathcal{N})$. Kosaki shows that $L_q(\mathcal{N})$ is a Banach dual of $L_p(\mathcal{N})$ for $\frac{1}{p} + \frac{1}{q} = 1$ with $p \in [1, \infty[$, and $L_\infty(\mathcal{N}) := \mathcal{N}$. The space $L_2(\mathcal{N})$ is a Hilbert space. $L_1(\mathcal{N})$ is isometrically isomorphic to $\mathcal{N}_\star$.

Non-commutative integration: Falcone&Takesaki'01

- Completion of canonical theory of integration on arbitrary W*-algebras $\mathcal{N}$.
- The relationship between Kosaki'80 and Falcone–Takesaki'01 is analogous to Dixmier'52 vs Segal'53 constructions of noncommutative L_p spaces with respect to f.n.s. traces: the former defined the spaces abstractly, as completions of vector spaces with respect to Banach norms, while the latter defined the spaces concretely, by the families of unbounded operators satisfying additional (quite nontrivial!) properties, allowing to define (very nontrivial) notion of noncommutative integral $\int$.
- Functorially associated full range of non-commutative L_p spaces $\mathcal{N} \mapsto L_p(\mathcal{N})$
- $L_1(\mathcal{N}) \cong \mathcal{N}_*$, $L_\infty(\mathcal{N}) \cong \mathcal{N}$, $L_1(\mathcal{N})^\star \cong L_\infty(\mathcal{N})$.
- The generic elements of $L_p(\mathcal{N})$ have the form $x\phi^{1/p}$.
- $L_2(\mathcal{N})$ can be naturally equipped with the Hilbert space structure

$$L_2(\mathcal{N}) \times L_2(\mathcal{N}) \ni (x, y) \mapsto \langle x, y \rangle_{L_2(\mathcal{N})} := \int y^* x \in \mathbb{C}$$

- there is a bilinear Banach space duality pairing for $\frac{1}{p} + \frac{1}{q} = 1$, $p \in]1, \infty[$,

$$L_p(\mathcal{N}) \times L_q(\mathcal{N}) \ni (x, y) \mapsto [\![x, y]\!]_{\widetilde{\mathcal{N}}} := \int xy \in \mathbb{C}$$

- For any two f.n.s. weights on $\mathcal{N}$, $\psi, \phi \in \mathcal{W}_0(\mathcal{N})$, the equivalence relation $(x, \psi) \sim_t (y, \phi) \iff y = x[\psi : \phi]_t$ defines Banach spaces $L_{\mathrm{i}/t}(\mathcal{N}) := (\mathcal{N} \times \mathcal{W}_0(\mathcal{N}))/ \sim_t$, with elements denoted by $x\phi^{\mathrm{i}t}$.

Absolute integration theories in historical context

measure and integral equivalent:

- Integration on $\mathbb{R}^n$: Borel'1898, Lebesgue'19(01,04,10), Young'19(04,09,10), Radon'1913
- Integration theories on abstract commutative (function) spaces:
  1. abstract measure theory on countably additive algebras of subsets
     Fréchet'1915, Sierpiński'27'28, Nikodým'30, Kolmogorov'33, Maharam'42, Segal'51
  2. abstract measure theory on countably additive boolean algebras
     Carathéodory'38, Wecken'40, Loomis'47, Sikorski'48
  3. abstract integral theory on vector lattices
     Young'1911, Daniell'1919,20'21, Riesz'28'40, Kakutani'41, Stone'48'49

measure and integral are not equivalent, integration theory is strictly more general:

- Integration theories on noncommutative (operator) spaces:
  1. type I and II_1 W^*-algebras
     von Neumann–Murray'36–'43, von Neumann–Schatten'46'47'50
  2. semi-finite W^*-algebras with fixed n.s.f. trace
     Dye'52, Segal'53, Dixmier'53, Ogasawara–Yoshinaga'55, Kunze'58, Stinespring'59, Ovchinnikov'70'71, Nelson'74, Yeadon'73'75'80, Muratov'78'79, Fack–Kosaki'86, Dodds–Dodds–de Pagter'89'93, Sukochev–Chilin'90, Kunze'90, Kalton–Sukochev'08,...
  3. arbitrary W^*-algebras with fixed n.s.f. weight
     Haagerup'79–Terp'81, Connes'80–Hilsum'81, Masuda'83, Terp'82, Zolotarëv'83'85'88, Labuschagne'14,...
  4. arbitrary W^*-algebras (without a choice of a fixed weight)
     Woronowicz'79, Kosaki'80, Yamagami'92, Falcone'00, Falcone–Takesaki'01
- Integration theory on semi-finite nonassociative JBW-algebras:
  Ayupov'79–'86, Berdikulov'82'86, King'83, Abdullaev'83'84, Iochum'84'86, Haagerup–Hanche-Olsen'84, Trunov'85, Tadzhibaev'85'86, Ayupov–Abdullaev'85'90,...

Key takeout messages

- Main question: what are the points of state spaces? Pascal–Fermat '1654: probabilities. Huygens '1657: expectations. Integration theory on function spaces (1914–1951): both approaches (based on, respectively, measure and integral) are equivalent.
- Key fact: The setting of n.s.f. weights on W*-algebras allows to develop a full-fledged integration theory, which generalises wide range of objects and theorems of integration theory on measure spaces/vector lattices (e.g. partial integration, conditional expectations, $L_p(\mathcal{N})$ spaces, Orlicz spaces, etc...).
- Key fact #2: The noncommutative measure theory, focused on measures on the orthomodular lattices of projection operators (in any W*-algebra) is not equivalent, and has essentially less structure (e.g. it does not even allow to construct noncommutative L_p spaces).
- Hence, in noncommutative case Huygens wins with Fermat–Pascal: expectation/integral is more fundamental than probability measure.
- Key fact #3: Non-(type I) W*-algebras are indispensable generalisation of the setting of ordinary quantum mechanics in several important cases, e.g., to define the generalisation of maximum entropy states in thermodynamical limit (known as Kubo–Martin–Schwinger states), which are (in turn) required for the exact mathematical derivations of Fulling–Unruh and Hawking effects [Sewell'80, Sewell'82, Fredenhagen–Haag'90].

Objects = quant. information models = sets of quantum integrals

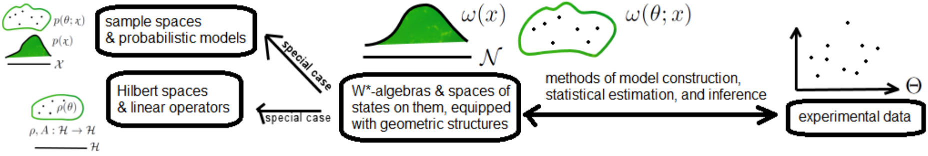
For any W*-algebra $\mathcal{N}$, $\mathcal{M}(\mathcal{N})$ will be defined as an arbitrary subset of a positive part of a Banach predual space of $\mathcal{N}$, $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_\star^+$.

Special cases:

- $\mathcal{N}$ is commutative $\Rightarrow \mathcal{M}(\mathcal{N}) = \mathcal{M}(\mathcal{X}, \mu)$
- $\mathcal{N}$ is type I $\Rightarrow \mathcal{M}(\mathcal{N}) = \mathcal{M}(\mathcal{H})$.

We do not assume that:

- $\mathcal{M}(\mathcal{N})$ is convex ($\iff$ probabilistic mixing)
- $\mathcal{M}(\mathcal{N})$ is smooth ($\iff$ asymptotic estimation)
- $\mathcal{M}(\mathcal{N})$ is normalised ($\iff$ frequentist interpretation)
- elements of $\mathcal{M}(\mathcal{N})$ are decomposable into tensor products ($\iff$ no initial correlations)



What are the morphisms?

- commutative probability theory:
  Bayes'1763–Laplace'1774 rule
  Kolmogorov'33: conditional expectations
  Wald'39: markovian (= normalised positive linear) maps
- quantum theory:
  von Neumann'32–Lüders'51 'projective state reduction' rule
  Umegaki'54: conditional expectations
  Stinespring'55: completely positive maps ("quantum markovian")

- all those mappings can be viewed as inductive inference, e.g. change state due to change of information

$T_\star$: markovian maps = coarse grainings

- Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be arbitrary W*-algebras.
- A function $T : \mathcal{N}_1 \to \mathcal{N}_2$ is called:
  1) positive iff $T(\mathcal{N}_1^+) \subseteq \mathcal{N}_2^+$;
  2) n-positive iff

$$T \otimes \mathrm{id}_{\mathrm{M}_n(\mathbb{C})} : \mathcal{N}_1 \otimes \mathrm{M}_n(\mathbb{C}) \ni x \otimes y \mapsto T(x) \otimes y \in \mathcal{N}_2 \otimes \mathrm{M}_n(\mathbb{C})$$

  is positive for $n \in \mathbb{N}$;
  3) completely positive (CP) iff it is n-positive $\forall n \in \mathbb{N}$ [Stinespring'55].
- In the commutative case, a coarse graining is defined as a positive linear function

$$T_\star : L_1(\mathcal{X}_1, \mho_1(\mathcal{X}_1), \mu_1) \to L_1(\mathcal{X}_2, \mho_2(\mathcal{X}_2), \mu_2)$$

  such that $\|f\| = \|T_\star(f)\|$ $\forall f \in L_1(\mathcal{X}_1, \mho_1(\mathcal{X}_1), \mu_1)^+$.
- In the noncommutative case, a coarse graining (CPTP) is defined as a positive linear function

$$T_\star : (\mathcal{N}_2)_\star \to (\mathcal{N}_1)_\star$$

  such that:
  1) There exists a completely positive map $T : \mathcal{N}_1 \to \mathcal{N}_2$ such that

$$(T_\star(\phi))(x) = \phi(T(x)) \;\; \forall \phi \in \mathcal{N}_2 \;\; \forall x \in \mathcal{N}_1$$

  2) $\|T_\star(\phi)\| = \|\phi\|$ $\forall \phi \in \mathcal{N}_2^+$ ($\iff$ $T(\mathbb{I}) = \mathbb{I}$)

Quantum maps: why complete positivity?

- According to Stinespring'55 theorem, trace preserving (e.g. $\mathrm{tr}(T(\rho)) = \mathrm{tr}(\rho)$) completely positive maps are characterised among all positive maps $T : \mathfrak{G}_1(\mathcal{H})^+_1 \to \mathfrak{G}_1(\mathcal{H})^+_1$ by the condition $\forall\rho \in \mathfrak{G}_1(\mathcal{H})^+_1\ \exists\mathcal{H}_{\mathrm{env}}\ \exists\rho_{\mathrm{env}} \in \mathfrak{G}_1(\mathcal{H}_{\mathrm{env}})^+_1\ \exists$ unitary U on $\mathcal{H}\otimes\mathcal{H}_{\mathrm{env}}$ s.t.

$$T(\rho) = \mathrm{tr}_{\mathcal{H}_{\mathrm{env}}}\left(U(\rho\otimes\rho_{\mathrm{env}})U^*\right).$$

So, one is restricted to CP(TP) maps iff one subscribes to the following paradigms:

1. All quantum evolutions arise from unitary evolutions
2. A nonunitary evolution arises iff:
   - it is possible to specify the Hilbert space and a quantum state of the "environment",
   - the quantum state subjected to unitary evolution is a tensor product of "system" and "environment" quantum states.

But:

1. do we always need to postulate a global unitary evolution?
2. do we always have a situation that the initial state of a system is noncorrelated/disentangled from an initial state of enviroment? (Reeh–Schlieder'61 theorem: this is generally never true for the vacuum state in (algebraic) QFT!)

Other reasons why CP maps are not necessary: Pechukas'94'95, Shaji&Sudarshan'04, ...

Quantum information theory: a summary of departure point

- positive trace class operators $\mathcal{T}(\mathcal{H})^+ := \{\rho \in \mathfrak{B}(\mathcal{H}) \mid \rho \geq 0,\ \mathrm{tr}_{\mathcal{H}}|\rho| < \infty\}$ replaced by the positive cone $\mathcal{N}_\star^+$
- general state spaces: arbitrary sets of denormalised quantum states: $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_\star^+$
- usually one assumes *a priori* that the morphisms $\mathcal{M}_1(\mathcal{N}_1) \to \mathcal{M}_2(\mathcal{N}_2)$ should be given by some CPTP maps, however there are seriously limiting assumptions behind it
- our main motivation is to find a reasonable class of (quantum, postquantum) state spaces and morphisms between them which would not be linear and would not be CPTP, yet would provide a consistent description of suitable quantum information processing tasks $\Rightarrow$ approach based on relative entropies D instead of tensor products $\otimes$

Quantum information informations/relative negentropies

Quantum information $D : \mathcal{M}(\mathcal{H}) \times \mathcal{M}(\mathcal{H}) \to [0, \infty]$
s.t. $D(\rho, \sigma) = 0 \iff \rho = \sigma$.

E.g.

- $D_1(\rho, \sigma) := \mathrm{tr}_{\mathcal{H}}(\rho \log \rho - \rho \log \sigma)$ [Umegaki'61]
- $D_{1/2}(\rho, \sigma) := 2\left\|\sqrt{\rho} - \sqrt{\sigma}\right\|^2_{\mathfrak{G}_2(\mathcal{H})} = 4\mathrm{tr}_{\mathcal{H}}(\frac{1}{2}\rho + \frac{1}{2}\sigma - \sqrt{\rho}\sqrt{\sigma})$
  (Hilbert–Schmidt norm^2)
- $D_{L_1(\mathcal{N})}(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_{\mathcal{T}(\mathcal{H})} = \frac{1}{2}\mathrm{tr}_{\mathcal{H}}|\rho - \sigma|$ (L_1/trace norm)
- $D_\gamma(\rho, \sigma) := \frac{1}{\gamma(1-\gamma)}\mathrm{tr}_{\mathcal{H}}(\gamma\rho + (1-\gamma)\sigma - \rho^\gamma\sigma^{1-\gamma})$; $\gamma \in \mathbb{R} \setminus \{0, 1\}$
  [Hasegawa'93]
- $D_{\chi^2}(\rho, \sigma) := \mathrm{tr}_{\mathcal{H}}((\sigma - \rho)\sigma^{-1}(\sigma - \rho))$ (quantum χ^2)
- $D_{\alpha,z}(\rho, \sigma) := \frac{1}{1-\alpha} \log \mathrm{tr}_{\mathcal{H}}(\rho^{\alpha/z}\sigma^{(1-\alpha)/z})^z$; $\alpha, z \in \mathbb{R}$
  [Audenauert–Datta'14]
- $D_{\mathfrak{f}}(\rho, \sigma) := \mathrm{tr}_{\mathcal{H}}(\sqrt{\rho}\,\mathfrak{f}(\mathfrak{L}_\rho\mathfrak{R}_\sigma^{-1})\sqrt{\rho})$; $\mathfrak{f}$ operator convex, $\mathfrak{f}(1) = 0$
  [Kosaki'82, Petz'85'86]

for $\mathrm{ran}(\rho) \subseteq \mathrm{ran}(\sigma)$, and with all $D(\rho, \sigma) := +\infty$ otherwise.

$D_{\mathfrak{f}}$: Quantum informations nonexpansive under coarse grainings

- A function $\mathfrak{f} : \mathbb{R}^+ \to \mathbb{R}$ is called operator convex [Kraus'36] iff
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \;\; \forall x, y \in \mathfrak{B}(\mathcal{H})^+ \;\; \forall t \in [0,1].$$
- A class of quantum informations [Kosaki'82,Petz'85] $D_{\mathfrak{f}} : \mathcal{N}_\star^+ \times \mathcal{N}_\star^+ \to [0, +\infty]$ s.t.
$$D_{\mathfrak{f}}(\omega, \phi) := \begin{cases} \langle \zeta(\phi), \mathfrak{f}(\Delta_{\omega,\phi})\zeta(\phi)\rangle_{L_2(\mathcal{N})} & : \omega \ll \phi \\ +\infty & : \text{ otherwise,} \end{cases}$$
for operator convex $\mathfrak{f}$ with $\mathfrak{f}(0) \leq 0$ and $\mathfrak{f}(1) = 0$, was shown by Petz'85 to satisfy
$$D_{\mathfrak{f}}(\rho, \sigma) \geq D_{\mathfrak{f}}(T_\star(\rho), T_\star(\sigma)) \; \forall \rho, \sigma \in \mathcal{N}_\star^+ \; \forall T_\star \text{ s.t. } \mathrm{dom}(T_\star) = \mathcal{N}_\star^+ \qquad (2)$$
for $\mathfrak{f}$ bounded from above, and any 2-positive (hence also any CPTP) T, with $=$ attained iff T is a normal $*$-isomorphism.
- Tomamichel–Colbeck–Renner'09: proof of (**??**) for type I $\mathcal{N}$ without boundedness assumption, and CPTP T. Hiai–Mosonyi–Petz–Bény'11: (**??**) with no boundedness, for type I $\mathcal{N}$ and any Schwarz (hence 2-positive, hence CPTP) T.
- In particular, for $\mathfrak{f}(t) = t\log(t) - (t-1)$, this gives Araki'76'77 information (its nonexpansivity for any Schwarz T was proved by Uhlmann'77)
$$D_1(\omega, \phi) = \begin{cases} (\phi - \omega)(\mathbb{I}) + \langle \zeta(\omega), \log(\Delta_{\omega,\phi})\zeta(\omega)\rangle_{L_2(\mathcal{N})} & : \; \omega \ll \phi \\ +\infty & : \text{ otherwise.} \end{cases}$$
- Hence, $D_1(\omega, \phi)$ of normalised states is an $L_2(\mathcal{N})$-expectation value of a relative modular hamiltonian. (Another curious relationship, after every faithful state being Gibbs state w.r.t. modular automorphism.)

$D_\mathfrak{f}$: Properties, special cases and commutative analogue

- Characterisation of $D_\mathfrak{f}$ by means of nonexpansivity under $T_\star$ is not known.
- $D_\mathfrak{f}(\omega,\phi)$ is lower semi-continuous on $\mathcal{N}_\star^+ \times \mathcal{N}_{\star 0}^+$ with product of norm topologies.
- If $\mathfrak{f}(0)=0$ then $D_\mathfrak{f}(\omega,\phi)$ is jointly convex in ω and ϕ.
- For $\mathcal{N}=\mathfrak{B}(\mathcal{H})$: $D_\mathfrak{f}(\rho,\sigma) := \mathrm{tr}_\mathcal{H}(\rho^{1/2}\mathfrak{f}(\mathfrak{L}_\rho\mathfrak{R}_\sigma^{-1})\rho^{1/2})$ if $\mathrm{ran}(\rho)\subseteq\mathrm{ran}(\sigma)$ and $=+\infty$ otherwise [Petz'86]. E.g.:
  - $\mathfrak{f}(\lambda)=\lambda\log\lambda \Rightarrow D_\mathfrak{f}(\rho,\sigma)=\mathrm{tr}_\mathcal{H}(\rho\log\rho-\rho\log\sigma)$ [Umegaki'61; monot.: Lindblad'74]
  - $\mathfrak{f}(\lambda)=(\lambda-1)^2 \Rightarrow D_\mathfrak{f}(\rho,\sigma)=\mathrm{tr}_\mathcal{H}((\sigma-\rho)\sigma^{-1}(\sigma-\rho))$ (quantum χ^2)
- For a commutative $\mathcal{N}$ Kosaki–Petz $D_\mathfrak{f}$ becomes a special case of:
- Let $\mathfrak{f}:\mathbb{R}^+\to\mathbb{R}$ be convex on $]0,\infty[$ with $\mathfrak{f}(1)=0$, strictly convex at 1, and $\mathfrak{f}(0):=\lim_{\lambda\to+0}\mathfrak{f}(\lambda)>-\infty$. Let $\mu_\omega \ll \upsilon \gg \nu_\phi$. Then the Csiszár'63–Morimoto'63 $\mathfrak{f}$-information is defined as $D_\mathfrak{f}: L_1(\mathcal{X},\upsilon)^+\times L_1(\mathcal{X},\upsilon)^+\to[0,\infty]$ s.t.

$$D_\mathfrak{f}(\omega,\phi) := \int \nu_\phi \mathfrak{f}\left(\frac{\mu_\omega}{\nu_\phi}\right) \text{ for } \mu_\omega \ll \nu_\phi \text{ and } +\infty \text{ otherwise.}$$

- E.g.: Pearson'1900–Kagan'63 χ^2-distance: $\mathfrak{f}(\lambda)=(\lambda-1)^2 \Rightarrow \int\nu_\phi\left(\frac{\mu_\omega}{\nu_\phi}-1\right)^2$, Hellinger'1909–Kakutani'48 distance: $\mathfrak{f}(\lambda)=(1-\sqrt{\lambda})^2 \Rightarrow \int(\sqrt{\mu_\omega}-\sqrt{\nu_\phi})^2$, Kullback–Leibler'51 information: $\mathfrak{f}(\lambda)=\lambda\log(\lambda)\Rightarrow D_1(\omega,\phi)=\int\mu_\omega\log\frac{\mu_\omega}{\nu_\phi}$.
- Csiszár'78 characterised $D_\mathfrak{f}$ for finite sample spaces by the conditions of: 1) nonexpansivity under coarse grainings $T_\star$, 2) invariance for Radon–Nikodým quotient invariance under $T_\star$, 3) joint convexity, 4) additive decomposability of $T_\star$ under all exclusive–and–exhaustive partitions of sample space.

Duality of $D_{\mathfrak{f}}$

- Csiszár'75:

$$\mathfrak{f}^{\mathbf{c}}(\lambda) := \lambda \mathfrak{f}(\frac{1}{\lambda}) \quad \text{for} \quad \lambda > 0$$

$$\mathfrak{f}^{\mathbf{c}}(0) := \lim_{\lambda \to \infty} \frac{1}{\lambda} \mathfrak{f}(\lambda) \in]-\infty, +\infty],$$

then $\mathfrak{f}^{\mathbf{c}} : \mathbb{R}^+ \to \mathbb{R}$ is convex on $]0, \infty[$, and $\mathfrak{f}^{\mathbf{cc}} = \mathfrak{f}$.

- Vajda'72:

$$D_{\mathfrak{f}}(\omega, \phi) = D_{\mathfrak{f}^{\mathbf{c}}}(\phi, \omega) \iff \exists t \in \mathbb{R} \ \forall \lambda \in]0, \infty[\ \mathfrak{f}(\lambda) - \mathfrak{f}^{\mathbf{c}}(\lambda) = (\lambda - 1)t,$$

$$D_{\mathfrak{f}}(\omega, \phi) = D_{\mathfrak{f}}(\phi, \omega) \Leftarrow \mathfrak{f}(\lambda) = \mathfrak{f}^{\mathbf{c}}(\lambda) \ \forall \lambda \in]0, \infty[\, .$$

- For example,

$$\mathfrak{f}_\gamma(t) = \begin{cases} \frac{1}{\gamma} + \frac{1}{1-\gamma} t - \frac{1}{\gamma(1-\gamma)} t^\gamma & : \gamma \in]0, 1[\\ t \log t - (t-1) & : \gamma = 1 \\ -\log t + (t-1) & : \gamma = 0, \end{cases}$$

$$\mathfrak{f}^{\mathbf{c}}_\gamma(t) = \begin{cases} \frac{1}{\gamma(1-\gamma)}(1 - t^{1-\gamma}) + \frac{1}{\gamma}(t-1) & : \gamma \in]0, 1[\\ t \log t - (t-1) & : \gamma = 0 \\ -\log t + (t-1) & : \gamma = 1. \end{cases}$$

$$\mathfrak{f}_0(t) = \lim_{\gamma \to^+ 0} \mathfrak{f}_\gamma(t) = \mathfrak{f}^{\mathbf{c}}_1(t), \quad \mathfrak{f}_1(t) = \lim_{\gamma \to^- 1} \mathfrak{f}_\gamma(t) = \mathfrak{f}^{\mathbf{c}}_0(t).$$

D_γ

This gives [Jenčová'03–Ojima'03]:

$$D_{\mathfrak{f}_\gamma}(\omega,\phi) = \left\langle \zeta(\phi), \left(\frac{1}{\gamma} + \frac{1}{1-\gamma}\Delta_{\omega,\phi} - \frac{1}{\gamma(1-\gamma)}\Delta^{\gamma}_{\omega,\phi}\right)\zeta(\phi)\right\rangle_{L_2(\mathcal{N})}$$

In particular [Hasegawa'93]: $D_\gamma(\rho,\sigma) := \frac{1}{\gamma(1-\gamma)}\mathrm{tr}_{\mathcal{H}}(\gamma\rho + (1-\gamma)\sigma - \rho^\gamma\sigma^{1-\gamma})$ for $\gamma \in \mathbb{R}\setminus\{0,1\}$

as well as [Araki'76'77]:

$$\begin{aligned} D_0(\omega,\phi) &= \left\langle \zeta(\phi), \left(-\log(\Delta_{\omega,\phi}) + \Delta_{\omega,\phi} - \mathbb{I}\right)\zeta(\phi)\right\rangle_{L_2(\mathcal{N})} \\ &= (\omega-\phi)(\mathbb{I}) - \left\langle \zeta(\phi), \log(\Delta_{\omega,\phi})\zeta(\phi)\right\rangle_{L_2(\mathcal{N})}, \\ D_1(\omega,\phi) &= \left\langle \zeta(\phi), \left(\Delta_{\omega,\phi}\log(\Delta_{\omega,\phi}) - \Delta_{\omega,\phi} + \mathbb{I}\right)\zeta(\phi)\right\rangle_{L_2(\mathcal{N})} \\ &= (\phi-\omega)(\mathbb{I}) + \left\langle \zeta(\phi), \left(\Delta_{\omega,\phi}\log(\Delta_{\omega,\phi})\right)\zeta(\phi)\right\rangle_{L_2(\mathcal{N})} \\ &= (\phi-\omega)(\mathbb{I}) + \left\langle \zeta(\omega), \log(\Delta_{\omega,\phi})\zeta(\omega)\right\rangle_{L_2(\mathcal{N})}. \end{aligned}$$

Hence, $\phi \ll \omega \ll \phi \;\Rightarrow\; D_\gamma(\omega,\phi) = D_{1-\gamma}(\phi,\omega) \quad \forall\gamma \in [0,1]$,

$$D_\gamma(\omega,\phi) = D_\gamma(\phi,\omega) \iff \gamma = \frac{1}{2},$$

$$\begin{aligned} D_{1/2}(\phi,\psi) &= 2(\phi+\psi)(\mathbb{I}) - 4\left\langle \zeta(\phi), \Delta^{1/2}_{\psi,\phi}\zeta(\phi)\right\rangle_{L_2(\mathcal{N})} = 2(\phi+\psi)(\mathbb{I}) - 4\left\langle \zeta(\phi), \zeta(\psi)\right\rangle_{L_2(\mathcal{N})} \\ &= 2\|\zeta(\phi) - \zeta(\psi)\|^2_{L_2(\mathcal{N})}, \end{aligned}$$

$$D_{1/2}(\rho,\sigma) := 2\left\|\sqrt{\rho} - \sqrt{\sigma}\right\|^2_{\mathfrak{G}_2(\mathcal{H})} = 4\mathrm{tr}_{\mathcal{H}}\left(\frac{1}{2}\rho + \frac{1}{2}\sigma - \sqrt{\rho}\sqrt{\sigma}\right).$$

RPK'11: Canonical form of D_γ

- Liese–Vajda'87, Zhu–Rohwer'95 (independent of representing measure):

$$D_\gamma(\omega,\phi) = \int \frac{\gamma\mu_\omega + (1-\gamma)\nu_\phi - \mu_\omega^\gamma \nu_\phi^{1-\gamma}}{\gamma(1-\gamma)},$$

- Jenčová'03–Ojima'03 (dependent on representing weight):

$$D_\gamma(\omega,\phi) = \frac{\gamma\omega(\mathbb{I}) + (1-\gamma)\phi(\mathbb{I}) - \mathrm{re}\left[\left[u_\omega \Delta^\gamma_{\omega,\psi}, u_\phi \Delta^{1-\gamma}_{\phi,\psi}\right]\right]_\psi}{\gamma(1-\gamma)},$$

- RPK'11: Using Falcone–Takesaki integral, two above formulations became unified into a general form:

$$D_\gamma(\omega,\phi) := \int \left(\frac{\gamma\omega + (1-\gamma)\phi - \omega^\gamma \phi^{1-\gamma}}{\gamma(1-\gamma)} \right)$$

and, including boundary terms, $\tilde{\gamma} \in [0,1]$:

$$D_{\tilde{\gamma}}(\omega,\phi) := \int \lim_{\gamma\to\tilde{\gamma}} \left(\frac{\gamma\omega + (1-\gamma)\phi - \omega^\gamma \phi^{1-\gamma}}{\gamma(1-\gamma)} \right).$$

- This includes Petz'85 (representing weight independent):

$$D_1(\phi,\omega) = i \lim_{t\to 0} \frac{\phi}{t} \left(\left(\frac{\mathrm{D}\omega}{\mathrm{D}\phi} \right)_t - \mathbb{I} \right).$$

Entropic paradigm: absolute and relative

- Gibbs'1902, Elsasser'37, Jaynes'57, Jaynes'62–Zubarev'62,...:

  constrained maximisation of **absolute** entropy
  (e.g., $S_{vN}(\rho) = -D_1(\rho, \psi) + \log \dim \mathcal{H}$, with a fixed prior $\psi = \mathbb{I}/\dim \mathcal{H}$)
  as a method of **model construction**:

  $$\rho(\text{constraints}) := \arg\sup\{S_{vN}(\omega) \mid \text{constraints}(\omega)\}$$

  selecting a specific class of models $\mathcal{M}$ with elements parametrised by allowed values of constraints' parameters and maximally noninformative with respective to anything else.

- Kullback'59, Good'63, Hobson'69,...:

  minimisation of $D_1(\rho, \psi)$ as a method of **state transformation (estimation, learning, updating,...)** from ψ into an element of the set that satisfies given constraints and is maximally noninformative with respective to anything else.

Max.Ent. approach to foundations of statistical mechanics I

Ordinary approaches (kinetic, ergodic) to foundations of stat. mech. are based on frequentist interpretation of probability. They are unable to deal with nonequilibrium stat. mech.

Elsasser'37–Jaynes'57 approach:

1. Use bayesian interpretation of probability
2. Specify constraints C describing your knowledge (theoretic assumptions and experimental data)
3. The predictive probability density p is determined by the maximum of Gibbs['1902]–Shannon['49] information entropy $\mathbf{S}_{\mathrm{GS}}$ under these constraints,

$$p := \arg\sup_{q \in C} \left\{ -\sum_j q(x_j) \log q(x_j) \right\} \equiv \arg\sup_{q \in C} \{\mathbf{S}_{\mathrm{GS}}(q)\}, \quad (3)$$

which exists uniquely if C is a closed and convex set of probabilities.

For example: assuming:

- discrete sample (=phase) space $\mathcal{X}$,
- some linearly independent functions ("observables") $\{f_i : \mathcal{X} \to \mathbb{R} \mid i \in \{1, \ldots, n\}\}$,
- constraints $C = \left\{\sum_j q(x_j) f_i(x_j) = \theta_i\right\}$, where $\theta_i \in \mathbb{R}$ can be defined e.g. as arithmetic averages of experimental data measured in experiment,

the solution of (**??**) reads $p(x) = \exp\left(-\sum_{i=1}^{n} \lambda_i f_i(x)\right)$, with Lagrange multipliers λ_i uniquely determined by the constraints.

Max.Ent. approach to foundations of statistical mechanics II

For nonequilibrium case:

- let $(\mathcal{X}, \mu)$ be a measure space,
- let $\alpha_t : \mathcal{X} \to \mathcal{X}$ be a family of μ-preserving automorphisms with $t \in \mathbb{R}$,
- let $\theta_1, \ldots, \theta_m$ be parameters with the ranges $\Theta_1, \ldots, \Theta_m$,
- let observables be given by $f_1(\chi, t), \ldots, f_m(\chi, t)$, with $f_k(\chi, t) := f_k(\alpha_t(\chi), 0)$
- let experimental data be provided by quantities $a_1(\theta_1, t), \ldots, a_m(\theta_m, t)$, which are incorporated by the constraints

$$C := \left\{ \int_{\mathcal{X}} \mu(\chi) q(\chi, \theta_k, t) f_k(\chi, t) = a_k(\theta_k, t) \right\}.$$

Then the solution of constrained entropy maximisation reads

$$q(\chi, t) = \exp\left(-\sum_{k=1}^{m} \int_{\Theta_k} \mathrm{d}\theta_k \int_{t_0}^{t} \mathrm{d}\tilde{t}\, \lambda_k(\chi, \theta_k, \tilde{t}) f_k(\chi, \tilde{t}) \right),$$

with Lagrange multipliers λ_k determined by the constraints.

For finite dimensional quantum case: use density matrices ρ, linearly independent self-adjoint operators f_i, and von Neumann entropy $-\rho \log \rho$. [Jaynes'62, Zubarev'62, equivalence proved in: Zubarev–Kalashnikov'70]

Segal entropy

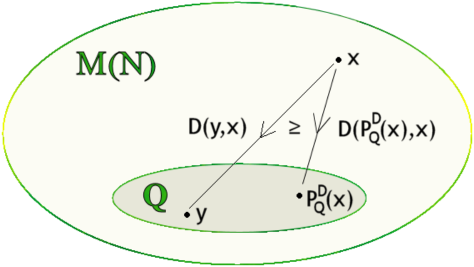
- The procedure of constrained maximum Gibbs–Shannon/von Neumann entropy works very well (e.g., characterises classical/quantum Gibbs states) for $\dim \mathcal{X} < \infty$ / $\dim \mathcal{H} < \infty$ (i.e., for commutative and noncommutative type $\mathrm{I}_{n<\infty}$ W^*-algebras).
- Segal'60 proposed a joint generalisation of both entropies, that includes type II W*-algebras: $\mathbf{S}_{\mathrm{Seg}}(\omega) := -\tau(\rho \log \rho)$, where ρ is a noncommutative Radon–Nikodým quotient of $\omega \in \mathcal{N}_\star^+$ w.r.t. f.n.s. trace τ on $\mathcal{N}$.
- This allowed him to extend maximum entropy characterisation of Gibbs states to include type II_1 W*-algebras.
- However, for types I_∞ and II_∞, for every $\phi \in \mathcal{N}_\star^+$, the open neighbourhood of ϕ w.r.t. $d_{\mathcal{N}_\star}(\phi, \omega) := \frac{1}{2}\|\phi - \omega\|_{\mathcal{N}_\star}$ (i.e., $L_1(\mathcal{N})$ norm distance) contains a dense set of $\mathbf{S}_{\mathrm{Seg}}(\omega) = +\infty$.
- (Yet, if $\rho := \mathrm{e}^{-\beta H}$ with $\tau(\rho) < \infty$, then $\tau(\sigma H) < \tau(\rho H) \Rightarrow \mathbf{S}_{\mathrm{Seg}}(\sigma) < \infty$.)
- Question 1: how to generalise max.ent./Gibbs/thermodynamic equilibrium states to continuous (type I_∞, II_∞, III) case? Answer: Use KMS states. Main idea: KMS condition generalises the property: If $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, $\dim \mathcal{H} < \infty$, $\alpha_t = \mathrm{e}^{\mathrm{i}th}(\cdot)\mathrm{e}^{-\mathrm{i}th}$, $t \in \mathbb{R}$, $h \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}}$, then $\omega = \mathrm{tr}_{\mathcal{H}}(\rho\,\cdot\,) \in \mathcal{N}_\star^+$ with $\rho = \mathrm{e}^{-\beta h}$ satisfies:
$$\mathrm{tr}_{\mathcal{H}}(\rho x y) = \mathrm{tr}_{\mathcal{H}}(\mathrm{e}^{-\beta h} x \mathrm{e}^{\beta h} \mathrm{e}^{-\beta h} y) = \mathrm{tr}(\mathrm{e}^{-\beta h} y \alpha_{\mathrm{i}\beta}(x)) = \mathrm{tr}(\rho y \alpha_{i\beta}(x)).$$
- Question 2: How to generalise max.ent. nonequilibrium states? Answer: Use constrained relative entropy maximisation, because it is well defined for any W*-algebras. (In finite-dimensional commutative case such approach to nonequilibrium thermodynamics was proposed and developed by Schlögl'66–75.)

Quantum entropic projections

Let $\mathcal{Q} \subseteq \mathcal{N}_\star^+$ be such that for each $\psi \in \mathcal{M}(\mathcal{N})$ there exists a unique solution

$$\mathfrak{P}^D_{\mathcal{Q}}(\psi) := \arg\inf_{\rho\in\mathcal{Q}} \{D(\rho,\psi)\}\,.$$

It will be called an entropic projection.



- for $D_{1/2}(\rho,\sigma) = 4\mathrm{tr}_{\mathcal{H}}(\frac{1}{2}\rho + \frac{1}{2}\sigma - \sqrt{\rho}\sqrt{\sigma})$, and $\mathcal{Q}$ defined as images of closed convex subsets $\tilde{\mathcal{Q}} \subseteq \mathfrak{G}_2(\mathcal{H})^+$ under the mapping $\tilde{\mathcal{Q}} \ni \sqrt{\rho} \mapsto \rho \in \mathcal{Q}$
- for $\tilde{\mathcal{Q}}$ given by the closed linear subspaces of the Hilbert–Schmidt (GNS) space $\mathfrak{G}_2(\mathcal{H})$, the entropic projections $\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}$ coincide with the ordinary linear projection operators on $\mathfrak{G}_2(\mathcal{H})$.
- for $D_1(\rho,\sigma) = \mathrm{tr}_{\mathcal{H}}(\rho\log\rho - \rho\log\sigma)$ and $\mathcal{M}(\mathcal{H}) = \mathcal{T}(\mathcal{H})^+_1$, $\psi \in \mathcal{T}(\mathcal{H})^+_1$, $h \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}}$, then [Araki'77, Donald'90]

$$\exists!\ \psi^h := \arg\inf_{\rho\in\mathcal{T}(\mathcal{H})^+_1} \{D_1(\rho,\psi) + \mathrm{tr}_{\mathcal{H}}(\rho h)\}\,.$$

  Here the codomains of $\mathfrak{P}^{D_1}_{\mathcal{Q}}$ are the hypersurfaces of the fixed expectation value of h, which is a direct generalisation and relativisation of maximum Gibbs–Shannon/von Neumann/Segal entropy principle to Umegaki/Araki D_1.
- All of the above holds for arbitary W*-algebra $\mathcal{N}$, not only $\mathfrak{B}(\mathcal{H})$.

Bayes–Laplace rule and maximum relative entropy

- Fundamental principle of statistical inference in the bayesian statistics:

  the Bayes'1763–Laplace'1774 rule: $$p(x) \mapsto p_{\mathrm{new}}(x) := p(x|b) = \frac{p(x)p(b|x)}{p(b)}.$$

- Williams'80, Warmuth'05, Caticha&Giffin'06: the Bayes–Laplace rule is a special case of
$$p(\chi) \mapsto p_{\mathrm{new}}(\chi) := \arg\inf_{q \in \mathcal{Q}} \left\{ D_1(q,p) \right\},$$
  where D_1 is the Kullback–Leibler information
$$D_1(q,p) := \int_{\mathcal{X}} \mu(\chi) q(\chi) \log\left(\frac{q(\chi)}{p(\chi)}\right).$$

- Douven&Romeijn'12: the Bayes–Laplace rule is also a special case of
$$p \mapsto \arg\inf_{q \in \mathcal{Q}} \left\{ D_1(p,q) \right\} = \mathfrak{P}^{D_0}_{\mathcal{Q}}(p),$$
  where $D_0(p,q) = D_1(q,p)$.

Caticha–Giffin'06'08 derivation

for $p, q \in \mathcal{M} := L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_1^+$, $\dim \mathcal{M} =: n < \infty$, with parametrisation $\theta : \mathcal{M} \to \Theta \subseteq \mathbb{R}^n$ allowing to consider a measure space $(\Theta, \mho_{\mathrm{Borel}}(\Theta), \mathrm{d}\theta)$ as well as a product measure space $(\mathcal{X} \times \Theta, \mho(\mathcal{X} \times \Theta), \tilde{\mu} \times \mathrm{d}\theta)$, consider a constrained minimisation of D_1:

$$p(x,\theta) \mapsto p_{\mathrm{new}}(x,\theta) := \arg\inf_{q(x,\theta)\in\mathcal{M}} \left\{ \int_{\mathcal{X}} \tilde{\mu}(x) q(x,\theta) \log\left(\frac{q(x,\theta)}{p(x,\theta)}\right) + F(q(x,\theta)) \right\}, \tag{4}$$

$$F(q(x,\theta)) = \lambda_1 \left(\int_{\mathcal{X}} \tilde{\mu}(x) \int_{\Theta} \mathrm{d}\theta q(x,\theta) - 1 \right) + \lambda_2(x) \left(\int_{\Theta} \mathrm{d}\theta q(x,\theta) - \boldsymbol{\delta}(x-b) \right), \tag{5}$$

where λ_1 and $\lambda_2(x)$ are Lagrange multipliers, and $\boldsymbol{\delta}(x-b)$ is Dirac's delta at $b \in \mathcal{X}$. The posterior probability selected as a unique solution of this variational problem is given by

$$p_{\mathrm{new}}(x,\theta) = \frac{p(x,\theta)\mathrm{e}^{\lambda_2(x)}}{\int_{\mathcal{X}} \tilde{\mu}(x) \int_{\Theta} \mathrm{d}\theta p(x,\theta)\mathrm{e}^{\lambda_2(x)}}, \tag{6}$$

where $\lambda_2(x)$ is determined via $\frac{\int_{\Theta} \mathrm{d}\theta p(x,\theta)\mathrm{e}^{\lambda_2(x)}}{\int_{\mathcal{X}} \tilde{\mu}(x) \int_{\Theta} \mathrm{d}\theta p(x,\theta)\mathrm{e}^{\lambda_2(x)}} = \boldsymbol{\delta}(x-b)$. Hence,

$$p_{\mathrm{new}}(x,\theta) = \frac{p(x,\theta)\boldsymbol{\delta}(x-b)}{\int_{\Theta} \mathrm{d}\theta p(x,\theta)} = \frac{p(x,\theta)\boldsymbol{\delta}(x-b)}{p(x)} =: \boldsymbol{\delta}(x-b) p(\theta|x), \tag{7}$$

which leads to the Bayes–Laplace rule on Θ,

$$p(\theta) \mapsto p_{\mathrm{new}}(\theta) = \int_{\mathcal{X}} \tilde{\mu}(x) \boldsymbol{\delta}(x-b) p(\theta|x) = p(\theta|b), \tag{8}$$

whenever μ is such that $\int_{\mathcal{X}} \tilde{\mu}(x)\boldsymbol{\delta}(x-b)h(x) = h(b)$ (for example, if $\tilde{\mu}(x) = \mathrm{d}x$).

Jeffrey's rule from maximum relative entropy

Caticha–Giffin'06'08:

If the second constraint in (**??**) is replaced by a more general form,

$$F(q(\chi,\theta)) = \lambda_1 \left(\int_{\mathcal{X}} \tilde{\mu}(\chi) \int_{\Theta} \mathrm{d}\theta q(\chi,\theta) - 1 \right) + \lambda_2(\chi) \left(\int_{\Theta} \mathrm{d}\theta q(\chi,\theta) - f(\chi) \right), \quad (9)$$

corresponding to a condition $q(\chi) = \int_{\Theta} \mathrm{d}\theta q(\chi,\theta) = f(\chi)$ with a given probability density $f \in \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, then the entropic projection (**??**) reproduces Jeffrey's rule on Θ,

$$p_{\mathrm{new}}(x,\theta) = \frac{p(\chi,\theta)}{p(\chi)} f(\chi) =: p(\chi|\theta) f(\chi) = p(\chi|\theta) p_{\mathrm{new}}(\chi), \quad (10)$$

$$p(\theta) \mapsto p_{\mathrm{new}}(\theta) = \int_{\mathcal{X}} \tilde{\mu}(\chi) f(\chi) \frac{p(\chi,\theta)}{p(\chi)} = \int_{\mathcal{X}} \tilde{\mu}(\chi) p(\theta|\chi) f(\chi) = \int_{\mathcal{X}} \tilde{\mu}(\chi) p(\theta|\chi) p_{\mathrm{new}}(\chi). \quad (11)$$

Lüders' rules

- Lüders' rules [von Neumann'32, Lüders'51] provide the basic paradigm for the description of quantum state change due to measurement of an observable $x = \sum_i \lambda_i P_i$:

$$\rho \mapsto \rho_{\mathrm{new}} := \sum_i P_i \rho P_i \;\; (\text{'weak'} = \text{'nonselective'}),$$

$$\rho \mapsto \rho_{\mathrm{new}} := \frac{P \rho P}{\mathrm{tr}_{\mathcal{H}}(P\rho)} \;\; (\text{'strong'} = \text{'selective'})$$

- Bub'77'79, Caves–Fuchs–Schack'01, Fuchs'02, Jacobs'02: Lüders' rules should be considered as rules of inference (conditioning) that are quantum analogues of the Bayes–Laplace rule.
- Yet, no mathematically exact relationship was provided.

Quantum bayesian inference from quantum entropic projections

- F.Hellmann–W.Kamiński–RPK'14:

  1. weak Lüders' rule is a special case of $\rho \mapsto \arg\inf_{\sigma\in\mathcal{Q}} \{D_1(\rho,\sigma)\}$ with
  $$\mathcal{Q} = \{\sigma \in \mathcal{T}(\mathcal{H})^+ \mid [P_i,\sigma] = 0\ \forall i\}$$
  2. strong Lüders' rule derived from $\rho \mapsto \arg\inf_{\sigma\in\mathcal{Q}} \{D_1(\rho,\sigma)\}$ with
  $$\mathcal{Q} = \{\sigma \in \mathcal{T}(\mathcal{H})^+ \mid [P_i,\sigma] = 0,\ \mathrm{tr}_{\mathcal{H}}(\sigma P_i) = p_i\ \forall i\}$$
  under the limit $p_2,\ldots,p_n \to 0$.
  3. hence, weak and strong Lüders' rules are special cases of quantum entropic projection $\mathfrak{P}^{D_0}_{\mathcal{Q}}$ based on relative entropy $D_0(\sigma,\rho) = D_1(\rho,\sigma)$.

Bayes–Laplace and Lüders' conditionings are special cases of entropic projections
$\Rightarrow$ "quantum bayesianism $\subseteq$ quantum relative entropism".

Quantum Jeffrey's rule

- Jeffrey'65: proposed another rule for probabilistic bayesian inference, generalising the Bayes–Laplace rule:

$$p(x|\eta) \mapsto p_{\mathrm{new}}(x|\eta) := \sum_{i=1}^{n} p(x|b_i)\lambda_i = \sum_{i=1}^{n} \frac{p(x \wedge b_i|\eta)}{p(b_i|\eta)}\lambda_i,$$

  where $n \in \mathbb{N}$,
  - $\{b_1, \ldots, b_n\}$ is a set of exhaustive and mutually exclusive elements of boolean algebra,
  - $\lambda_i = p_{\mathrm{new}}(b_i|\eta)\ \forall i \in \{1, \ldots, n\}$,
  - $p(b_i|\eta) \neq 0$.
- Caticha&Giffin'06: under more general constraints $\mathcal{Q}$, one can derive Jeffrey's rule as a special case of $\mathfrak{P}^{D_1}_{\mathcal{Q}}$
- F.Hellmann–W.Kamiński–RPK'14: derivation of a quantum analogue of Jeffrey's rule:

$$\mathcal{T}(\mathcal{H})^+_1 \ni \rho \mapsto \rho_{\mathrm{new}} := \arg\inf_{\sigma \in \mathcal{Q}} \{D_1(\rho, \sigma)\} = \sum_{i=1}^{n} \frac{P_i \rho P_i}{\mathrm{tr}_{\mathcal{H}}(\rho P_i)}\lambda_i \in \mathcal{T}(\mathcal{H})^+_1,$$

  where $n \in \mathbb{N}$,
  - $\{P_1, \ldots, P_n\} \subseteq \mathrm{Proj}(\mathfrak{B}(\mathcal{H}))$, $\sum_{i=1}^{n} P_i = \mathbb{I}$, $P_i P_j = \boldsymbol{\delta}_{ij} P_i$,
  - $\lambda_i = \mathrm{tr}_{\mathcal{H}}(\rho_{\mathrm{new}} P_i)\ \forall i \in \{1, \ldots, n\}$,
  - $\mathrm{tr}_{\mathcal{H}}(\rho P_i) \neq 0$.

  It generalises Lüders' rule.

Quantum measurements from quantum entropic projections

- Hence: the rule of maximisation of relative entropy (entropic projection on the subset determined by constraints) can be considered as a nonlinear generalisation of the dynamics describing *elementary* "quantum measurement".
- F.Hellmann–W.Kamiński–RPK'14: also quantum analogue of Jeffreys' rule follows
- M.Munk-Nielsen'15: partial trace is also entropic projection (at least for strictly positive states)
- more measurements and more general results: RPK–M.Munk-Nielsen'20 (under construction)
- these results are for D_0 and/or D_1; however there are many more Ds...

Earlier results (obtained exclusively for symmetric information functionals):

- Herbut'69: weak Lüders' rule is a special case of $\mathfrak{P}^{d_2}_{\mathcal{Q}}$, with $d_2(\rho,\sigma) = \left\langle \sqrt{\rho} - \sqrt{\sigma}, \sqrt{\rho} - \sqrt{\sigma} \right\rangle_{\mathfrak{G}_2(\mathcal{H})}$.
- Hadjisavvas'81: strong von Neumann rule is a special case of $\mathfrak{P}^{d_1}_{\mathcal{Q}}$ with $d_1(\rho,\sigma) = \frac{1}{2}\mathrm{tr}_{\mathcal{H}}|\rho - \sigma|$.
- Raggio'84: strong von Neumann rule is a special case of maximum Cantoni–Uhlmann transition probability $\iff \mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}$.

Quantum entropic projections: towards general setting

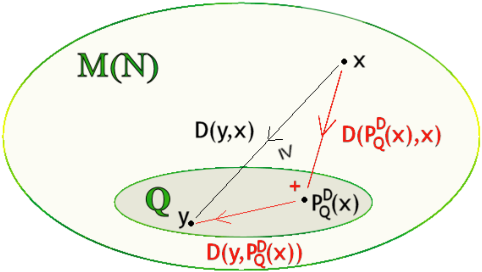
- How general class of nonlinear state transformations/updating ("due to measurement/information gain") can be derived from entropic projections, allowing both D and Q to vary?
- The choice of the class of sets $\mathcal{Q}$ for which $\exists!\mathfrak{P}^D_{\mathcal{Q}}$ depends very strongly on the structure of D (and vice versa!): the choice of discrimination functional (D) defining the principle of inference ($\mathfrak{P}^D_{\mathcal{Q}}$) determines the accepted data types ($\mathcal{Q}$), and conversely.
- Considering $\mathcal{Q}$ as objects and $\mathfrak{P}^D_{\mathcal{Q}}$ as candidates for morphisms, this leads to a question of general conditions on families $(\mathcal{Q}, D)$ guaranteeing the existence and uniqueness of $\mathfrak{P}^D_{\mathcal{Q}}$, together with good composition properties of subsequent projections $\Rightarrow$ the general problem is to define (and eventually characterise) the categories of entropic projections.
- In analogy to (nonexpansivity under coarse grainings determining the structure of $D_{\mathfrak{f}}$), we need some principle constraining $D/\mathcal{Q}$ that would guarantee existence, uniqueness, and good composition properties of D-projections.
- This principle will be provided by: (1) the generalised pythagorean inequality, which will be equivalently expressed in terms of (2) the local–to–global property of convex functions, and in terms of (3) the Young–Fenchel inequality for the Legendre case of Fenchel duality.

Generalised pythagorean inequality/equation

- We say that D satisfies a generalised pythagorean inequality at $\mathcal{Q}$ iff [Brègman'67–Chencov'68]

$$D(\phi,\psi) \geq D(\phi,\mathfrak{P}^D_{\mathcal{Q}}(\psi)) + D(\mathfrak{P}^D_{\mathcal{Q}}(\psi),\psi) \;\; \forall(\phi,\psi) \in \mathcal{Q} \times \mathcal{M}.$$

- In particular, in the case of equality, information decomposes additively under a projection onto a suitable subspace, hence we have a nonlinear, yet additive (!), decomposition: **data = signal + noise**



- Goal: introduce the class of relative negentropies D, for which
  1. $\exists!\mathfrak{P}^D_{\mathcal{Q}}$ iff Q is convex and closed (in a suitable sense!)
  2. generalised pythagorean inequality always holds
  3. generalised pythagorean equality holds iff Q is affine (in a suitable sense!)

Generalised pythagorean equation: examples

- Example 1: If $\mathcal{Q}$ forms an affine subset of $L_1(\mathcal{X},\mu)^+$, then [Brègman'67–Chencov'68], and D_1 is the Kullback–Leibler information, then
$$D_1(\phi,\psi^h)+D_1(\psi^h,\psi)=D_1(\phi,\psi)\;\;\forall(\phi,\psi)\in\mathcal{Q}\times L_1(\mathcal{X},\mu)^+.$$
- Example 2: If $\mathcal{Q}$ forms an affine subset of $\mathfrak{G}_2(\mathcal{H})$, then:
$$\left\|x-\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}(z)\right\|^2_{\mathfrak{G}_2(\mathcal{H})}+\left\|\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}}(z)-z\right\|^2_{\mathfrak{G}_2(\mathcal{H})}=\|x-z\|^2_{\mathfrak{G}_2(\mathcal{H})}.$$
- Example 3: If $\mathcal{Q}:=\{\phi\in\mathcal{N}^+_\star\mid\phi(\mathbb{I})=1,\;\;\phi(h)=\mathrm{const}\}$, i.e., if it is an affine subset of $\mathcal{N}^+_{\star1}$, and D_1 is the normalised Araki information, then [Donald'90]
$$D_1(\phi,\psi^h)+D_1(\psi^h,\psi)=D_1(\phi,\psi)\;\;\forall(\phi,\psi)\in\mathcal{Q}\times\mathcal{N}^+_{\star1}.$$
- Observation: Convexity and affinity in these examples are defined w.r.t. to different linear structure (Ex.1: $L_1(\mathcal{X},\mu)$ space, Ex.2: $L_2(\mathcal{N})$ space, Ex.3: $L_1(\mathcal{N})$ space).

Convexity: local vs global

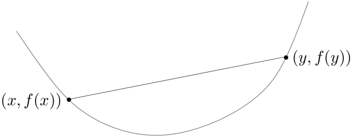
- A function $f : C \to \mathbb{R}$, with convex $C \subseteq \mathbb{R}^n$, is called convex iff

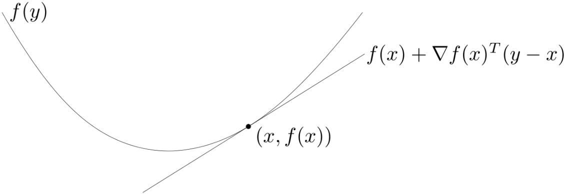
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \;\; \forall x, y \in C \;\; \forall \lambda \in [0,1].$$

- If $C \subseteq \mathbb{R}^n$ is convex and open, and if $f : C \to \mathbb{R}$ is differentiable, then f is convex iff

$$f(y) \geq f(x) + [\nabla f(x)]^\top (y - x) \;\; \forall x, y \in C, \tag{12}$$

where $\nabla \equiv \mathrm{grad}$.





- r.h.s. of (**??**) = first order Taylor approximation of f near x = supporting hyperplane through $(x, f(x))$ = linear witness.
- In other words, r.h.s. of (**??**) is a local approximation of ["information about"] f which is a global underestimator of ["information about"] f.
- Boyd S., Vandenberghe L., *Convex optimization* (2004): «This is perhaps the most important property of convex functions (...) and convex optimization problems».
- E.g., $\nabla f(x) = 0 \Rightarrow f(y) \geq f(x) \; \forall y \in C$, i.e. x is a global minimum of f.

Brègman'67: $\widetilde{D}_\Psi$: first idea

Let $\Psi : \mathbb{R}^n \to]-\infty, \infty]$ be proper (i.e., $\mathrm{efd}(\Psi) := \{x \in \mathbb{R}^n \mid \Psi(x) \neq \infty\} \neq \varnothing$), strictly convex and differentiable on $\mathrm{int}(\mathrm{efd}(\Psi))$. Then $\widetilde{D}_\Psi : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty]$,

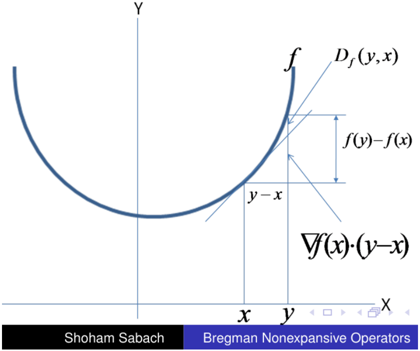
$$\widetilde{D}_\Psi(y,x) := \begin{cases} \Psi(y) - \Psi(x) - \sum_{i=1}^n (y-x)_i [(\nabla\Psi)(x)]^i & : x \in \mathrm{int}(\mathrm{efd}(\Psi)) \\ +\infty & : \text{otherwise} \end{cases}$$

Properties:

- $\widetilde{D}_\Psi(y,x)$ is convex in y
- $\widetilde{D}_\Psi(y,x) \geq 0$, with $=0$ iff $x = y$
- $\widetilde{D}_{\Psi+\lambda\Phi} = \widetilde{D}_\Psi + \lambda \widetilde{D}_\Phi$ for $\lambda \geq 0$
- in general: $\widetilde{D}_\Psi(y,x) \neq \widetilde{D}_\Psi(x,y)$
- given a convex closed $Q \subseteq \mathrm{int}(\mathrm{efd}(\Psi))$,

$$\widetilde{D}_\Psi(y,x) \geq \widetilde{D}_\Psi(y, \mathfrak{P}_Q^{\widetilde{D}_\Psi}(x)) + \widetilde{D}_\Psi(\mathfrak{P}_Q^{\widetilde{D}_\Psi}(x), x),$$

  with equality iff Q is affine ($\iff$ generalised pythagorean equation).



Al'ber–Butnariu'97, Butnariu–Iusem'00,...: $\widetilde{D}_\Psi$ is characterised by the generalised pythagorean inequality (for projections onto closed convex sets), or, equivalently, by generalised pythagorean equality (for projections onto closed affine sets).

$\widetilde{D}_\Psi$: examples

- $X = (\mathbb{R}^n)^+$, $\Psi(x) = \sum_{i=1}^n x_i \log(x_i)$, $\nabla\Psi(x) = \log(x)$,
$$\widetilde{D}_\Psi(x,y) = \sum_{i=1}^n (y_i - x_i + x_i(\log(x_i) - \log(y_i))) = D_1(x,y)$$
- $X = \mathcal{H}$ (Hilbert space with $\dim\mathcal{H} < \infty$), $\Psi(x) = \frac{1}{2}\|x\|^2_{\mathcal{H}}$, $\nabla\Psi = \mathrm{id}_{\mathcal{H}}$,
$$\widetilde{D}_\Psi(x,y) = \frac{1}{2}\|x-y\|^2_{\mathcal{H}} = 4\widetilde{D}_{1/2}(x,y)$$
- $X =]0,\infty[^n$, $\Psi = -\sum_{i=1}^n \log(x_i)$ (Burg'67'75), $\nabla\Psi(x) = -\frac{1}{x}$, then [Pinsker'60/Itakura–Saito'68]:
$$\widetilde{D}_\Psi(x,y) = \sum_{i=1}^n \left(-\log\frac{x_i}{y_i} + \frac{x_i}{y_i} - 1\right) \quad \forall (x,y) \in]0,\infty[^{2n}.$$

$\widetilde{D}_\Psi$: towards $\dim X = \infty$

How to generalise to $\dim X = \infty$?

- X will be assumed to be a Banach space, while the gradient ∇ will be generalised to Gateaux derivative.
- If X is a topological vector space over $\mathbb{K}$, $t \in \mathbb{R}$, and $\Psi : X \to]-\infty, +\infty]$ is proper then the Gateaux'1914 derivative of Ψ at $x \in X$ in the direction $h \in X$ reads

$$X \times X \ni (x, h) \mapsto \mathfrak{D}^{\mathrm{G}}\Psi(x; h) := \lim_{t \to 0} \frac{\Psi(x + th) - \Psi(x)}{t} \in [-\infty, \infty]. \quad (13)$$

- If x is fixed and (**??**) exists for all $h \in X$, and is (linear and bounded) in h, then Ψ is called Gateaux differentiable at x.
- If X is a Banach space and Ψ is Gateaux differentiable at $x \in X$, then

$$\mathfrak{D}^{\mathrm{G}}\Psi(x; y) =: [\![y, \mathfrak{D}^{\mathrm{G}}_x\Psi]\!]_{X \times X^\star} \quad \forall y \in X \quad (14)$$

  defines the Gateaux derivative $\mathfrak{D}^{\mathrm{G}}_x\Psi$.
- Mazur'33: Given a Banach space X with a unit ball $X_{\leq 1}$, $\|\cdot\|_X$ is Gateaux differentiable at $x \in X \setminus \{0\}$ iff $\|x\|_X \cdot X_{\leq 1}$ has a unique supporting hyperplane at x.

$\widetilde{D}_\Psi$: reflexive case [Bauschke–Borwein–Combettes'01]

Let X be a reflexive Banach space ($X \cong X^{\star\star}$), let $\Psi : X \to]-\infty, \infty]$ be Legendre (= convex, proper, lower semi-continuous, Gateaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ with some additional conditions [see next slide]). Then, $\widetilde{D}_\Psi : X \times X \to [0, \infty]$,

$$\widetilde{D}_\Psi(x,y) := \begin{cases} \Psi(x) - \Psi(y) - [\![x-y, \mathfrak{D}^{\mathrm{G}}_y\Psi]\!]_{X\times X^\star} & : y \in \mathrm{int}(\mathrm{efd}(\Psi)) \\ +\infty & : \text{otherwise} \end{cases} \tag{15}$$

satisfies:

- $\widetilde{D}_\Psi(x,y) = 0 \iff x = y$ (information)
- $\widetilde{D}_\Psi(x,y) + \widetilde{D}_\Psi(y,z) = \widetilde{D}_\Psi(x,z) + [\![x-y, \mathfrak{D}^{\mathrm{G}}_z\Psi - \mathfrak{D}^{\mathrm{G}}_y\Psi]\!]_{X\times X^\star}$ (generalised cosine theorem)
- if $C \subseteq X$ is convex and closed then

$$\forall y \in \mathrm{int}(\mathrm{efd}(\Psi)) \quad \exists! \mathfrak{P}^{\widetilde{D}_\Psi}_C(y) := \arg\inf_{x\in C} \left\{ \widetilde{D}_\Psi(x,y) \right\}$$

- if C is furthermore also an affine subset of X then

$$\widetilde{D}_\Psi(x, \mathfrak{P}^{\widetilde{D}_\Psi}_C(y)) + \widetilde{D}_\Psi(\mathfrak{P}^{\widetilde{D}_\Psi}_C(y), y) = \widetilde{D}_\Psi(x,y) \quad \forall (x,y) \in C \times X$$

- If $X = \mathcal{H}$, $\Psi_{1/2} = \frac{1}{2}\|\cdot\|^2_{\mathcal{H}}$, then $\mathfrak{D}^{\mathrm{G}}\Psi_{1/2} = \mathrm{id}_{\mathcal{H}}$, $\widetilde{D}_{\Psi_{1/2}}(x,y) = \frac{1}{2}\|x-y\|^2_{\mathcal{H}}$
- If X = reflexive Banach space with $\|\cdot\|_X$ Gateaux differentiable at unit sphere X_1, $\Psi = \frac{1}{2}\|\cdot\|^2_X$, then $\mathfrak{D}^{\mathrm{G}}\Psi = j : X \to X^\star$, i.e. a duality map $j(x) := \{y \in X^\star \mid y(x) = \|y\|_{X^\star}\|x\|_X,\ \|y\|_{X^\star} = \|x\|_X\} = \{*\}$.

Fenchel duality, subdifferential, and Legendre functions

- Given a proper function $\Psi : X \to [-\infty, \infty]$, and a duality pairing $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!] : X \times X^{\mathbf{d}} \to \mathbb{K})$, the Fenchel'49 dual $\Psi^{\mathbf{F}} : X^{\mathbf{d}} \to [-\infty, \infty]$ reads
$$\Psi^{\mathbf{F}}(\hat{y}) := \sup_{x \in X}\{\mathrm{re}\, [\![x, \hat{y}]\!]_{X \times X^{\mathbf{d}}} - \Psi(x)\} \;\; \forall \hat{y} \in X^{\mathbf{d}}.$$
- $\mathrm{efd}(\Psi) \neq \varnothing \Rightarrow \Psi^{\mathbf{F}}(\hat{y}) > -\infty \; \forall \hat{y} \in X^{\mathbf{d}}$
- $\Psi^{\mathbf{F}}$ and $\Psi^{\mathbf{FF}}$ are always convex.
- $\Psi^{\mathbf{FF}}|_X = \Psi$ if $(X, X^{\mathbf{d}})$ are dual pair of locally convex topological vector spaces with weak-$\star$ and weak- topologies, respectively, and Ψ is weakly lower semi-continuous and convex.
- The Fenchel'49 subdifferential of a proper $\Psi : X \to]-\infty, \infty]$ at $x \in \mathrm{efd}(\Psi)$ is
$$\partial\Psi(x) := \{\hat{y} \in X^{\mathbf{d}} \mid \Psi(z) - \Psi(x) \geq \mathrm{re}\, [\![z - x, \hat{y}]\!]_{X \times X^{\mathbf{d}}} \forall z \in X\}. \qquad (16)$$
For $x \in X \setminus \mathrm{efd}(\Psi)$, $\partial\Psi(x) := \varnothing$.
- If X is a reflexive Banach space, then a proper, convex, lower semi-continuous $\Psi : X \to]-\infty, \infty]$ is called Legendre iff $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$, $\mathrm{int}(\mathrm{efd}(\Psi^{\mathbf{F}})) \neq \varnothing$, $\partial\Psi$ is single valued on $\mathrm{efd}(\partial\Psi) := \{x \in \mathrm{efd}(\Psi) \mid \partial\Psi(x) \neq \varnothing\}$, and $\partial\Psi^{\mathbf{F}}$ is single valued on $\mathrm{efd}(\partial\Psi^{\mathbf{F}}) := \{x \in \mathrm{efd}(\Psi^{\mathbf{F}}) \mid \partial\Psi^{\mathbf{F}}(x) \neq \varnothing\}$ [Rockafellar'67: $\mathbb{R}^n$ case, Bauschke–Borwein–Combettes'01: reflexive Banach space case].

Brègman functional from Young–Fenchel inequality [B–B–C'01]

- If $\Psi : X \to]-\infty, \infty]$ is convex and $\hat{y} \in X^{\mathsf{d}}$, then the Young–Fenchel inequality holds
$$\Psi(x) - \Psi^{\mathsf{F}}(\hat{y}) - \mathrm{re}\,[\![x, \hat{y}]\!] \geq 0,$$
with equality iff $\hat{y} \in \partial\Psi(x)$.
- If X is a Banach space, and Ψ is proper, convex, lower semi-continuous, and Gateaux-differentiable at $x \in X$, then $\partial\Psi(x) = \{\mathfrak{D}^{\mathrm{G}}_x\Psi\}$.
- If Ψ is Legendre, then it is (Gateaux differentiable and strictly convex) on $\mathrm{int}(\mathrm{efd}(\Psi)) = \mathrm{efd}(\mathfrak{D}^{\mathrm{G}}\Psi)$ and Ψ^{F} is (Gateaux differentiable and strictly convex) on $\mathrm{int}(\mathrm{efd}(\Psi^{\mathsf{F}})) = \mathrm{efd}(\mathfrak{D}^{\mathrm{G}}\Psi^{\mathsf{F}})$.
- If Ψ is Legendre, then
$$\mathfrak{D}^{\mathrm{G}}\Psi : \mathrm{int}(\mathrm{efd}(\Psi)) \to \mathrm{int}(\mathrm{efd}(\Psi^{\mathsf{F}}))$$
is a bijection, with $\mathfrak{D}^{\mathrm{G}}\Psi^{-1} = \mathfrak{D}^{\mathrm{G}}\Psi^{\mathsf{F}}$.
- So, for Legendre Ψ, the one-sided Fenchel duality becomes a two-sided Legendre duality.
- Two key consequences:
$$\widetilde{D}_\Psi(x,y) = \Psi(x) - \Psi(y) - [\![x - y, \mathfrak{D}^{\mathrm{G}}_y\Psi]\!] = \Psi(x) + \Psi^{\mathsf{F}}(\mathfrak{D}^{\mathrm{G}}_y\Psi) - [\![x, \mathfrak{D}^{\mathrm{G}}_y\Psi]\!]$$
$$\widetilde{D}_\Psi(x,y) = \widetilde{D}_{\Psi^{\mathsf{F}}}(\mathfrak{D}^{\mathrm{G}}_y\Psi, \mathfrak{D}^{\mathrm{G}}_x\Psi).$$
- Hence, the Brègman functional $\widetilde{D}_\Psi(x,y)$ can be seen as an information functional characterising the content of the Young–Fenchel inequality in the Legendre case of Fenchel duality.

D_Ψ: Postquantum Brègman informations [RPK'17]

- Let X be a reflexive Banach space, $\Psi : X \to]-\infty,\infty]$ a Legendre function, let U be (a subset of) a positive generating cone of a base norm space Y, let $\ell : U \to \ell(U) \subseteq \mathrm{int}(\mathrm{efd}(\Psi)) \subseteq X$ be a bijection. We define a postquantum Brègman information as:

$$D_\Psi(\phi,\omega) := \widetilde{D}_\Psi(\ell(\phi),\ell(\omega))$$

  where $\widetilde{D}_\Psi$ is a Brègman functional on X.
- The bijectivity of ℓ allows to induce a topology from X onto U.
- The existence and uniqueness of the projections onto $\mathcal{Q} \subseteq U$ is guaranteed by requiring $\ell(\mathcal{Q})$ to be convex and closed.
- One can think of ℓ as a (nonlinear) coordinate system on U, and X as the linear parameter space used for specification of the data required for the entropic projection.
- As a result, all postquantum Brègman informations D_Ψ satisfy generalised pythagorean inequality/equality for sets $\mathcal{Q}$ that are (closed and convex/affine) under ℓ-embeddings.
- If Y is given by a self-adjoint part of a predual of W*-algebra, then D_Ψ is a quantum Brègman information.

Quantum Brègman informations: D_γ as example

- Jenčová'03/'05 (very inspiring paper): $U = \mathcal{N}_\star^+$ for a W*-algebra $\mathcal{N}$, $X = L_{1/\gamma}(\mathcal{N}, \psi)$: noncommutative L_p space w.r.t. f.n.s. weight ψ on $\mathcal{N}$, $p = \frac{1}{\gamma} \in]1, \infty[$, $\ell_\gamma(\phi) = \frac{1}{\gamma}\Delta^\gamma_{\phi,\psi}$, $\Psi_\gamma(x) = \frac{1}{1-\gamma}\|\gamma x\|^{1/\gamma}$

$$D_\gamma(\omega,\phi) = \frac{1}{1-\gamma}\omega(\mathbb{I}) + \frac{1}{\gamma}\phi(\mathbb{I}) + \frac{1}{\gamma(1-\gamma)}\left[\left[\Delta^\gamma_{\omega,\psi}, \Delta^{1-\gamma}_{\omega,\psi}\right]\right]_\psi \qquad (17)$$

- For $\gamma \in]0,1[$ one has [Jenčová'05]: $\exists! \, \mathfrak{P}^{D_\gamma}_{\mathcal{Q}}(\psi) := \arg\inf_{\phi\in\mathcal{Q}}\{D_\gamma(\phi,\psi)\}$:

$$D_\gamma(\omega,\psi) \geq D_\gamma(\omega, \mathfrak{P}^{D_\gamma}_{\mathcal{Q}}(\psi)) + D_\gamma(\mathfrak{P}^{D_\gamma}_{\mathcal{Q}}(\psi), \psi) \quad \forall(\omega,\psi) \in \mathcal{Q} \times \mathcal{N}_\star^+$$

  if the following conditions are satisfied:
  1) $\mathcal{Q}$ is nonempty,
  2) $\ell_\gamma(\mathcal{Q}) \subseteq L_{1/\gamma}(\mathcal{N}, \psi)$ is convex,
  3) $\mathcal{Q}$ is closed in the topology induced on $\mathcal{N}_\star^+$ by ℓ_γ^{-1} from the weak topology of $L_{1/\gamma}(\mathcal{N}, \psi)$.
- The proof of above theorem as given in Jenčová'05 does not use the theory of Brègman functionals on Banach spaces, however the Brègman functional structure of $D_\gamma = D_{\Psi_\gamma} \circ (\ell_\gamma, \ell_\gamma)$ is discussed there explicitly.
- Weak and norm closures of convex sets coincide for reflexive Banach spaces.

Quantum Brègman informations: D_γ as example (II)

- RPK'11/'13/'17: given $\gamma \in [0,1]$, $X = L_{1/\gamma}(\mathcal{N})$, $\ell_\gamma(\phi) = \frac{1}{\gamma}\phi^\gamma$, $\Psi_\gamma(x) = \frac{1}{1-\gamma}\|\gamma x\|^{1/\gamma}$, $\widetilde{D}_{\Psi_\gamma}(\ell_\gamma(\cdot), \ell_\gamma(\cdot))$ gives $\mathcal{N}^+_\star \times \mathcal{N}^+_\star \ni (\omega, \phi) \mapsto D_\gamma(\omega,\phi) \in [0,\infty]$ s.t.

$$\begin{cases} \int \frac{1}{\gamma(1-\gamma)}\left(\gamma\omega + (1-\gamma)\phi - \omega^\gamma \phi^{1-\gamma}\right) & : \gamma \in]0,1[,\ \omega \ll \phi \\ \int \lim_{\tilde{\gamma}\to^{\pm}\gamma} \frac{1}{\tilde{\gamma}(1-\tilde{\gamma})}\left(\tilde{\gamma}\omega + (1-\tilde{\gamma})\phi - \omega^{\tilde{\gamma}}\phi^{1-\tilde{\gamma}}\right) & : \gamma \in \{0,1\},\ \omega \ll \phi \\ +\infty & : \text{otherwise}, \end{cases}$$

  with $\tilde{\gamma} \to^+ \gamma$ for $\gamma = 0$ and $\tilde{\gamma} \to^- \gamma$ for $\gamma = 1$.
- $D_\Psi \cap D_\mathfrak{f} = D_\gamma$: characterisation, in finite dimensional case, under some conditions:
  - commutative: Amari'09 ($\gamma \in \mathbb{R}$),
  - quantum: RPK'13 (conjecture) '19 (proof) ($\gamma \in [-1,2]$).

  Under further restriction to $\phi(\mathbb{I}) = 1$, the characterised class restricts to $\{D_0, D_1\}$:
  - proved by Csiszár'91 in commutative case
  - remarked (without proof) by Petz'07 in noncommutative case.

Quantum Brègman informations: D_γ as example (III)

$$D_\gamma(\omega,\phi) = \int \frac{1}{\gamma(1-\gamma)}\left(\gamma\omega + (1-\gamma)\phi - \omega^\gamma\phi^{1-\gamma}\right)$$

$$\widetilde{\ell}_\gamma : \mathcal{N}_\star \ni \phi = |\phi|(\,\cdot\, u_\phi) \mapsto \tfrac{1}{\gamma} u_\phi |\phi|^\gamma \in L_{1/\gamma}(\mathcal{N}) \text{ is a bijection}$$

$$\Psi_\gamma : L_{1/\gamma}(\mathcal{N}) \ni x \mapsto \Psi_\gamma(x) := \tfrac{1}{1-\gamma}\int(\gamma x)^{1/\gamma} = \tfrac{1}{1-\gamma}\|\gamma x\|_{1/\gamma}^{1/\gamma}$$

$$\|\cdot\|_{1/\gamma} : \phi^\gamma \mapsto \|\phi^\gamma\|_{1/\gamma} := (|\phi|(\mathbb{I}))^\gamma = \|\phi\|_{\mathcal{N}_\star}^\gamma$$

$$\mathfrak{D}^{\mathrm{G}}_x\|\cdot\|_{1/\gamma}(y) = \|x\|_{1/\gamma}^{-1}\mathrm{re}\,\left[\!\left[y, j_{1/\gamma}(x)\right]\!\right],$$

$$\text{because} \quad j = (f^2)' = 2f\cdot f' \quad\Rightarrow\quad f' = \tfrac{1}{2f}\cdot j,$$

$$(\mathfrak{D}^{\mathrm{G}}_x\Psi_\gamma)(y) = \left(\mathfrak{D}^{\mathrm{G}}\left(\tfrac{1}{1-\gamma}\|\gamma x\|_{1/\gamma}^{1/\gamma}\right)\right)(y) = \left(\tfrac{1}{1-\gamma}\|\gamma x\|_{1/\gamma}^{1/\gamma-1}\mathfrak{D}^{\mathrm{G}}\|x\|_{1/\gamma}\right)(y)$$

$$\mathfrak{D}^{\mathrm{G}}\Psi_\gamma = \widetilde{\ell}_{1-\gamma}\circ\widetilde{\ell}_\gamma^{-1} : L_{1/\gamma}(\mathcal{N}) \ni \tfrac{1}{\gamma}u_\phi|\phi|^\gamma \mapsto \tfrac{1}{1-\gamma}u_\phi|\phi|^{1-\gamma} \in L_{1/(1-\gamma)}(\mathcal{N})$$

$$\Psi_\gamma^{\mathbf{F}} = \Psi_{1-\gamma}$$

$$\widetilde{D}_{\Psi_\gamma}(x,y) = \Psi_\gamma(x) + \Psi_{1-\gamma}(\mathfrak{D}^{\mathrm{G}}_y\Psi_\gamma) - \mathrm{re}\,\left[\!\left[x, \mathfrak{D}^{\mathrm{G}}_y\Psi_\gamma\right]\!\right]$$

$$D_{\Psi_\gamma}(\omega,\phi) = \widetilde{D}_{\Psi_\gamma}(\ell_\gamma(\omega),\ell_\gamma(\phi)) = \Psi_\gamma(\ell_\gamma(\omega)) + \Psi_{1-\gamma}(\ell_{1-\gamma}(\phi)) - \mathrm{re}\,[\![\ell_\gamma(\omega),\ell_{1-\gamma}(\phi)]\!]$$

$$= \tfrac{1}{1-\gamma}\left\|\gamma\tfrac{1}{\gamma}|\omega|^\gamma\right\|_{1/\gamma}^{1/\gamma} + \tfrac{1}{\gamma}\left\|(1-\gamma)\tfrac{1}{1-\gamma}|\phi|^{1-\gamma}\right\|_{1/(1-\gamma)}^{1/(1-\gamma)} - \tfrac{1}{\gamma(1-\gamma)}\left[\!\left[\omega^\gamma,\phi^{1-\gamma}\right]\!\right]$$

$$= \tfrac{1}{1-\gamma}\omega(\mathbb{I}) + \tfrac{1}{\gamma}\phi(\mathbb{I}) - \tfrac{1}{\gamma(1-\gamma)}\int\omega^\gamma\phi^{1-\gamma}.$$

D_Ψ: (Post)quantum Brègman informations: examples (IV)

(2) [RPK'19]: a generalisation of D_γ to nonassociative $L_{1/\gamma}(A,\tau)$ spaces over JBW-algebras A with f.n.s. trace τ (with explicit calculation of all properties naturally generalising from the $L_{1/\gamma}(\mathcal{N})$ case thanks to Iochum'84'86/Aupov'86 proof of uniform convexity and uniform Fréchet differentiability of $L_{1/\gamma}(A,\tau)$).

(3) [RPK'20]: a family D_Ψ over reflexive Orlicz ideals of self-adjoint compact operators $\mathfrak{G}_\Upsilon(\mathcal{H})^{\mathrm{sa}}$ over countably dimensional Hilbert spaces $\mathcal{H}$, with Υ given by an invertible Orlicz function such that both Υ and Υ^{Y} satisfy $\triangle_2$ condition, $\ell = \ell_\Upsilon : \mathfrak{G}_1(\mathcal{H})^+ \ni \rho \mapsto \Upsilon^{-1}(\rho) \in \mathfrak{G}_\Upsilon(\mathcal{H})^+$, and Ψ given by any spectral Legendre function $\Psi = f \circ \lambda$, where $f : l_\Upsilon \to]-\infty,\infty]$ is any rearrangement-invariant (i.e., symmetric) Legendre function on the Orlicz'36 sequence space l_Υ, while $\lambda : \mathfrak{G}_\Upsilon(\mathcal{H})^{\mathrm{sa}} \to l_\Upsilon$ is a spectral map introduced in Borwein–Read–Lewis–Zhu'99.

- for $\dim\mathcal{H} < \infty$ the map λ is a vector of eigenvalues, listed in nonincreasing order, and all setting has been developed by Lewis'96
- B–R–L–Z'99 consider only $\mathfrak{G}_{1/\gamma}(\mathcal{H})^{\mathrm{sa}}$ spaces, but all constructions directly apply to $\mathfrak{G}_\Upsilon(\mathcal{H})^{\mathrm{sa}}$,
- the proof of $f \circ \lambda$ being Legendre iff f is Legendre follows implicitly from the proof of the analogous statement for Gateaux differentiability in B–R–L–Z'99 (in $\dim\mathcal{H} < \infty$ case it has been proved explicitly in Lewis'96).

(4) more examples: later in this talk!

(5) Any base norm space Y with a weakly compact base (e.g., if dim $Y < \infty$, or if Y is a type I_2 JBW-factor) is reflexive, so then the construction of $\widetilde{D}_\Psi$ applies directly.

(6) Araki information D_1 is a quantum Brègman information only in the finite dimensional (Umegaki) case. In general, it is not associated naturally with any reflexive Banach space, however it is a limit of a family of quantum Brègman informations: $\lim_{\gamma\to^+1} D_\gamma(\omega,\phi) = D_1(\omega,\phi)$. It satisfies one-sided version of the generalised cosine theorem [Donald'90]. For commutative $\mathcal{N}$, it turns to Kullback–Leibler D_1, for which the one-sided (right) generalised cosine and pythagorean theorems in ∞-dim case were proved by [Chencov'68].

$D_{\mathfrak{f}}$ vs D_Ψ: different preferred morphisms $\iff$ different structure

Csiszár–Morimoto/Kosaki–Petz $D_{\mathfrak{f}}$:

- $\langle\zeta(\phi),(\cdot)\zeta(\phi)\rangle_{L_2(\mathcal{N})}$ is a bijective canonical cone representation $L_1(\mathcal{N})^+ \to L_2(\mathcal{N})^+$ of $\phi(\cdot)$
- $\mathfrak{f}:\mathbb{R}^+\to\mathbb{R}$ is operator convex, $\mathfrak{f}(0)\leq 0$ and $\mathfrak{f}(1)=0$, possibly bounded from above
- $D_{\mathfrak{f}}(\omega,\phi) := \langle\zeta(\phi),\mathfrak{f}(\Delta_{\omega,\phi})\zeta(\phi)\rangle_{L_2(\mathcal{N})}$
- $\mathfrak{f}\leftrightarrow\mathfrak{f}^{\mathbf{c}}$ duality implies $D_{\mathfrak{f}}(\omega,\phi)=D_{\mathfrak{f}^{\mathbf{c}}}(\phi,\omega)$
- $D_{\mathfrak{f}}(\rho,\sigma)\geq D_{\mathfrak{f}}(T_\star(\rho),T_\star(\sigma))\;\forall\rho,\sigma\in\mathcal{N}_\star^+\;\forall T_\star$ s.t. $\mathrm{dom}(T_\star)=\mathcal{N}_\star^+$, with $=$ when $T_\star$ is a $*$-isomorphism.

Quantum Brègman D_Ψ:

- $\ell:\mathcal{N}_\star^+\to\ell(\mathcal{N}_\star^+)\subseteq\mathrm{int}(\mathrm{efd}(\Psi))\subseteq X$ is a bijection into (the subset of the positive cone of) reflexive noncommutative Banach space X
- $\Psi: X\to]-\infty,\infty]$ is convex, proper, lower semi-continuous, Legendre
- $D_\Psi(\omega,\phi) := \Psi(\ell(\omega))-\Psi(\ell(\phi))-[[\ell(\omega)-\ell(\phi),\mathfrak{D}^{\mathrm{G}}\Psi(\ell(\phi))]]_{X\times X^\star}$
- $\Psi\leftrightarrow\Psi^{\mathbf{F}}$ duality implies $D_\Psi(\ell(\omega),\ell(\phi))=D_{\Psi^{\mathbf{F}}}(\mathfrak{D}^{\mathrm{G}}\Psi(\ell(\phi)),\mathfrak{D}^{\mathrm{G}}\Psi(\ell(\omega)))$
- $D_\Psi(\omega,\phi)\geq D_\Psi(\omega,\mathfrak{P}^{D_\Psi}_C(\phi))+D_\Psi(\mathfrak{P}^{D_\Psi}_C(\phi),\phi)\quad\forall(\omega,\phi)\in C\times\mathcal{N}_\star^+$, with $C\subseteq\mathcal{N}_\star^+$ ℓ-convex ℓ-closed, with $=$ when C is ℓ-affine.

Chencov's programme of categorical geometrostatistics (I)

- Hotelling'29 (unpublished), Rao'45, Jeffreys'46: independent discoveries that 'Fisher information matrix' is a riemannian metric tensor $\mathbf{g}^{\mathrm{FR}}$ on the space of strictly positive probabilities over finite dimensional sample space.
- Chencov'64: introduced an affine connection ($\nabla^{\gamma=0}$) on statistical manifold. (Acknowledges his wife, E.Morozova, for the suggestion; Dawid'75: independent rediscovery.)
- Chencov'65: paper "Categories of mathematical statistics" with subsets of positive cone of $L_1(\mathcal{X},\mu)$ spaces as objects, and coarse grainings as arrows (independently introduced in: Lawvere'62(unpublished) and Morse–Sacksteder'66).
- Chencov'68: generalised pythagorean theorem for Kullback–Leibler relative entropy in ∞-dim (independently: Brègman'67 for finite dim and any D_Ψ).
- Chencov'69: characterisation of all riemannian–affine geometries ($\mathbf{g}^{\mathrm{FR}}, \nabla^\gamma$) on spaces of probability densities on finite dimensional sample spaces that are nonexpansive under markovian morphisms. (Amari'80: Independent rediscovery of ∇^γ connections)
- Araki'74, Donald'90: generalised pythagorean theorem for Araki information.
- Ingarden–Janyszek–Kossakowski–Kawaguchi'82: The Taylor expansion of Umegaki D_1 gives Mori'55–Kubo'56–Bogolyubov'61 quantum riemannian metric.
- Eguchi'83'85: The Taylor expansion of Csiszár–Morimoto $D_\mathfrak{f}$ gives $\mathbf{g}^\mathfrak{f} = \mathbf{g}^{\mathrm{FR}}\mathfrak{f}''(1)$ while $\nabla^\mathfrak{f}$ coincide with ∇^γ with $1 - 2\gamma = 2\mathfrak{f}'''(1) + 3\mathfrak{f}''(1)$.
- Morozova–Chencov'85,'87,'89–Petz'96: characterisation of riemannian geometries of quantum state spaces that are nonexpansive under quantum markovian morphisms.
- Nagaoka'94'95–Hasegawa'95: ∇^0 and ∇^1 affine connections on quantum state spaces.
- Lesniewski–Ruskai'99: Taylor expansion of the Kosaki–Petz $D_\mathfrak{f}$ gives exactly the Morozova–Chencov–Petz metrics.
- Jenčová'03'04: characterisaton of the class $\nabla^\mathfrak{f}$ of quantum affine connections which are nonexpansive under quantum markovian morphisms, and of its dually flat subclass ∇^γ.

Chencov's programme of categorical geometrostatistics (II)

Chencov'72: monography summarising his '64–'72 work

(de facto: an extended version of his '69 habilitation thesis)

source of the photo:
www.keldysh.ru/memory/chentsov

On the first page of Introduction:

«The system of all statistical decision rules of all thinkable statistical problems taken together with a natural operation of composition forms an algebraic category. This category gives birth to a homogeneous geometry of families of probabilistic laws, in which the families play the role of 'figures', while decision laws describe 'movements'. Two families are congruent if and only if, when they are having the same statistical properties. The subject of this monography most exactly could be described by a notion 'geometrostatistics'.»

Система всех статистических решающих правил всех мыслимых статистических задач с естественной операцией композиции образует алгебраическую категорию. Эта категория порождает однородную геометрию семейств вероятностных законов, в которой семейства играют роль «фигур», а решающие правила описывают «движения». При этом два семейства конгруэнтны тогда и только тогда, когда они обладают эквивалентными статистическими свойствами.

Предмет настоящей монографии точнее всего было бы окрестить термином «геометростатистика».

Ченцов Н.Н., 1972, Статистические решающие правила и оптимальные выводы, Наука, Москва (Engl. transl. 1982, *Statistical decision rules and optimal inference*, American Mathematical Society, Providence)

Nonlinearity & convexity: outlook

Our main goal: construct the categories of nonlinear (quantum, postquantum) geometrostatistics, using Brègman relative entropies D_Ψ, their entropic projections, and even more general D_Ψ-well-behaving nonlinear morphisms*, together with the corresponding brègmannian geometry, instead of markovian (CPTP, positive linear) maps and corresponding $D_\mathfrak{f}$-geometries.
$\Rightarrow$ Nonlinear nonmarkovian version of "Chencov programme".

(*these morphisms will be given by ℓ-embeddings of so-called Brègman strongly quasi-nonexpansive maps)

Mottos:

- «the needed applications of global analysis to calculus of variations or continuum physics are usually nonlinear. Another unspoken presupposition of mainstream mathematics seems to be: *nonlinear is a generalization of linear and hence more difficult*. But there are important ways in which a nonlinear category can be simpler than the linear category of vector space objects in it.»

  F.W. Lawvere, 1998, *Volterra's functionals and covariant cohesion of space*

- «In fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity»

  R.T. Rockafellar, 1993, *Lagrange multipliers and optimality*

Categories of postquantum brègmannian entropic projections

The elementary setting:

- $\mathtt{Cvx}(\ell, \Psi)$:
  - objects: ℓ-closed ℓ-convex subsets of U (e.g., of a positive generating cone of a given base norm space), including $\varnothing$
  - morphisms: $\mathfrak{P}^{D_\Psi}_{Q}$, including $\varnothing$
  - composition: $\mathfrak{P}^{D_\Psi}_{Q_2} \circ \mathfrak{P}^{D_\Psi}_{Q_1} = \mathfrak{P}^{D_\Psi}_{Q_1 \cap Q_2}$

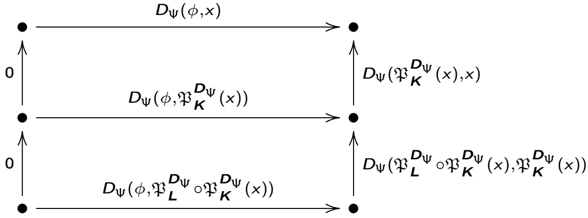
  Hence, it can be considered as the category of generalised pythagorean inequality.
- $\mathtt{Aff}(\ell, \Psi)$: as above, but Q restricted to ℓ-affine ℓ-closed sets: the category of generalised pythagorean equation
- $\mathtt{Cvx}^{\subseteq}(\ell, \Psi)$, $\mathtt{Aff}^{\subseteq}(\ell, \Psi)$: as two above, respectively, but with composition rule restricted to $Q_2 \subseteq Q_1$ (inclusion of convex/affine sets)
- Specific examples of above categories, with (ℓ, Ψ) determined by:
  1) spectral Legendre functions over Orlicz spaces of self-adjoint compact operators on countably dimensional Hilbert spaces
  2) noncommutative L_p spaces over arbitrary W*-algebras
  3) nonassociative L_p spaces over semi-finite JBW-algebras
  4) any base norm space Y with a weakly compact base and any Legendre function Ψ on Y
  5) more examples later in this talk

Brègman relative entropy as a functor (I)

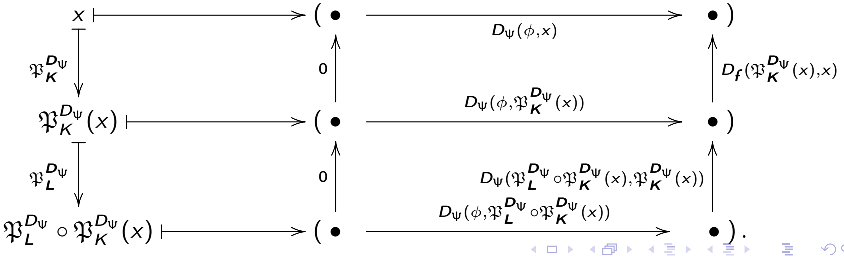
- Motivation: Baez–Fritz'14: characterisation of D_1 relative entropy as a functor from a suitable category into $[0, \infty]$.
- The class of Brègman relative entropies D_Ψ leads naturally to another functorial structure, arising from the generalised pythagorean theorem.
- $[0, \infty] :=$ a category consisting of one object $\bullet$, with morphisms given by the elements of the set $\mathbb{R}^+ \cup \{\infty\}$, and their composition defined by addition (Lawvere'73).
- $2 :=$ category consisting of two objects, one arrow between them, and the identity arrows on each of the objects.
- $[0, \infty]^2$ has objects given by morphisms of $[0, \infty]$, morphisms given by the commutative squares in $[0, \infty]$, and compositions given by commutative compositions of these squares.
- Let $\mathtt{Aff}^{\subseteq}_{Q}(\ell, \Psi)$ denote a full subcategory of $\mathtt{Aff}^{\subseteq}(\ell, \Psi)$, determined by the choice of its terminal object to be given by $Q \in \mathrm{Ob}(\mathtt{Aff}^{\subseteq}(\ell, \Psi))$.

Brègman relative entropy as a functor (II)

- Let $K_1, K_2, K_3, K, L \in \mathrm{Ob}(\mathtt{Aff}^{\subseteq}_{Q}(\ell,\Psi))$, $K \subseteq K_2$ and $L \subseteq K_3$.
- For each $\phi \in Q$, the generalised pythagorean theorem implies the commutativity of the diagram



which implies the commutativity of



Brègman relative entropy as a functor (III)

- This defines a contravariant functor $D_\Psi(\phi,\cdot):\mathtt{Aff}^{\subseteq}_{Q}(\ell,\Psi)\to[0,\infty]^2$.
- It naturally extends to the functor $D_\Psi(\phi,\cdot):\mathtt{Aff}^{\subseteq}(\ell,\Psi)\downarrow Q\to[0,\infty]^2$, where $\mathtt{Aff}^{\subseteq}(\ell,\Psi)\downarrow Q$ denotes a slice category of $\mathtt{Aff}^{\subseteq}(\ell,\Psi)$ over Q.
- For any two categories $\mathtt{C}$ and $\mathtt{D}$, the cartesian closedness of the category $\mathtt{Cat}$ of all small categories (with natural transformations as morphisms) implies that any functor $\mathtt{C}\to\mathtt{D}^2$ corresponds to a natural transformation in the functor category $\mathtt{D}^{\mathtt{C}}$.
- Hence, Q parametrises the family of natural transformations $D_\Psi(\phi,\cdot)$ in the category of functors $\mathtt{Aff}^{\subseteq}(\ell,\Psi)\downarrow Q\to[0,\infty]$.

Resource theoretic view

- Given any object Q in $\mathtt{Cvx}(\ell,\Psi)$, the set $\mathrm{Hom}_{\mathtt{Cvx}(\ell,\Psi)}(\,\cdot\,,Q)$ can be equipped with the structure of **commutative ordered monoid** via:
  - $\mathfrak{P}^{D_\Psi}_{Q_1}\wedge\mathfrak{P}^{D_\Psi}_{Q_2}:=\mathfrak{P}^{D_\Psi}_{Q_1\cap Q_2}$,
  - $\mathfrak{P}^{D_\Psi}_{Q_1}\leq\mathfrak{P}^{D_\Psi}_{Q_2}:=Q_1\subseteq Q_2$,
  - distinguished zero object given by $\mathfrak{P}^{D_\Psi}_{Q}$.
- Hence, each $\mathrm{Hom}_{\mathtt{Cvx}(\ell,\Psi)}(\,\cdot\,,Q)$ forms a resource theory in the sense of Fritz'17.
- Example: For $D_{1/2}$ defined by $X=$ Hilbert space $\mathcal{H}$, $\ell(\rho)=\rho^{1/2}$, $\Psi=\frac{1}{2}\|\cdot\|^2_{\mathcal{H}}$ and under restriction to such Q that correspond to closed linear subspaces of $\mathcal{H}$, the projections $\mathfrak{P}^{D_\Psi}_{Q}$ are given by the Hilbert space projection operators, while the operator implementing the finite join operation $\mathfrak{P}^{D_\Psi}_{Q_1}\wedge\ldots\wedge\mathfrak{P}^{D_\Psi}_{Q_n}$ is given by the von Neumann'33['50]–Kakutani'40–Halperin'62 theorem:

$$\lim_{k\to\infty}\left\|((P_{Q_n}\cdots P_{Q_1})^k-P_{Q_1\cap\ldots\cap Q_n})\xi\right\|_{\mathcal{H}}=0\;\;\forall\xi\in\mathcal{H}.$$

More on convergence of projectors

- von Neumann'33['50]–Kakutani'40 theorem: Let $Q_1, Q_2 \subset \mathcal{H}$ be closed subspaces of a Hilbert space $\mathcal{H}$ with $Q_1 \cap Q_2 \neq \varnothing$. Then $\mathrm{slim}_{k\to\infty}(P_{Q_1} \wedge P_{Q_2})^k = P_{Q_1\cap Q_2}$, i.e. $\lim_{k\to\infty} \left\|(P_{Q_1} P_{Q_2})^k - P_{Q_1\cap Q_2})\xi\right\|_{\mathcal{H}} = 0 \; \forall\xi \in \mathcal{H}$.
- Halperin'62: $\mathrm{slim}_{k\to\infty}(P_{Q_1} \wedge \ldots \wedge P_{Q_n})^k = P_{Q_1\cap\ldots\cap Q_n}$ for closed subspaces $Q_1, \ldots, Q_n$ of $\mathcal{H}$.
- P_Q in $\mathcal{H}$ is the same as the metric projection $\mathfrak{P}^{d_{\mathcal{H}}}_Q$, where $d_{\mathcal{H}}(x, y) := \|x - y\|_{\mathcal{H}}$, and it coincides with the Brègman projection $\mathfrak{P}^{D_\Psi}_Q$ for $\Psi : \mathcal{H} \to \mathbb{R}^+$ given by $\Psi(x) = \frac{1}{2}\|x\|^2_{\mathcal{H}}$.
- Brègman'65: If Q_1, Q_2 are closed and convex in $\mathcal{H}$, then the von Neumann–Kakutani algorithm converges weakly. (For finite dimensional $\mathcal{H}$ this implies norm convergence.)
- A mapping $T : \mathcal{D} \to \mathcal{H}$, $\mathcal{D} \subseteq \mathcal{H}$ is called nonexpansive iff $\|T(x) - T(y)\|_{\mathcal{H}} \leq \|x - y\|_{\mathcal{H}} \; \forall x, y \in \mathcal{D}$.
- A set of fixed points of T: $\mathrm{Fix}(T) := \{x \in \mathcal{D} \mid T(x) = x\}$.
- $\mathfrak{P}^{d_{\mathcal{H}}}_Q$ onto convex closed Q is nonexpansive with $\mathrm{Fix}(\mathfrak{P}^{d_{\mathcal{H}}}_Q) = Q$.

More on nonexpansivity

- In general Banach spaces X, $\mathfrak{P}^{dx}_{Q_1\cap\ldots\cap Q_n}$ may be ill-behaved, and the convergence of sequence of projections requires quite limiting assumptions.
- Two pathways: Brègman D_Ψ-projections ('67+) and nonexpansive operators ('65+).
- Brègman'67 [$\mathbb{R}^n$]: generate alternating sequence by $\mathfrak{P}^{D_\Psi}_{Q_n}\circ\ldots\circ\mathfrak{P}^{D_\Psi}_{Q_1}$; Theorem: it converges, under mild assumptions on D_Ψ.
- Browder'65–Göhde'65–Kirk'65: If K is bounded, closed, and convex subset of a uniformly convex Banach space X, and $T:K\to K$ is nonexpansive, then $\mathrm{Fix}(T)\neq\varnothing$.
- Bruck–Reich'77: Let X be a Banach space, then $T:\mathcal{D}\to X$ is called strongly nonexpansive iff it is nonexpansive (i.e., $\|T(x)-T(y)\|_X\leq\|x-y\|_X$ $\forall x,y\in\mathcal{D}$) and satisfies: if $(\{x_n-y_n\}_{n\in\mathbb{N}}$ is bounded and $\lim_{n\to\infty}(\|x_n-y_n\|_X-\|T(x_n)-T(y_n)\|_X)=0)$ then $\mathrm{slim}_{n\to\infty}((x_n-y_n)-(T(x_n)-T(y_n)))=0$.
- Bruck–Reich'77 theorem:
  1. Composition of strongly nonexpansive (SN) maps is SN.
  2. If T is SN and $\mathrm{Fix}(T)\neq\varnothing$ then $\mathrm{slim}_{n\to\infty}(T^n(x)-T^{n+1}(x))=0$.
  3. If X is uniformly convex, and $\{P_1,\ldots,P_k\}$ are norm-1 linear projections on X, then $\mathrm{slim}_{n\to\infty}(P_k\cdots P_1)^n=P$, where P is a norm-1 linear projection on X.

Brègman nonexpansive operators

- Let X be a Banach space $\Psi : X \to]-\infty, \infty]$ be proper, convex, lower semi-continuous, and Gateaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$. Let $\varnothing \neq M \subseteq \mathrm{int}(\mathrm{efd}(\Psi))$. Then $T : M \to \mathrm{int}(\mathrm{efd}(\Psi))$ will be called:
  - completely D_Ψ-nonexpansive (CN(Ψ)) iff $D_\Psi(T(x), T(y)) \leq D_\Psi(x, y)\ \forall x, y \in M$;
  - left strongly D_Ψ-quasi-nonexpansive (LSQ(Ψ)) iff [Censor–Reich'96, Reich'96]:

  1. $D_\Psi(x, T(y)) \leq D_\Psi(x, y)\ \forall (x, y) \in \widehat{\mathrm{Fix}}(T) \times M$,
  2. $(p \in \widehat{\mathrm{Fix}}(T)$, $\{x_n\}_{n\in\mathbb{N}}$ bounded, $\lim n \to \infty(D_\Psi(p, x_n) - D_\Psi(p, T(x_n))) = 0) \Rightarrow \lim_{n\to\infty} D_\Psi(T(x_n), x_n) = 0$,
  3. $\widehat{\mathrm{Fix}}(T) := \{x \in M \mid \exists$ a sequence $\{x_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \|x_n - T(x_n)\|_X = 0$ and $\{x_n\}_{n\in\mathbb{N}}$ is weakly convergent to $x\}$.

- Reich'96–Martín-Marquez–Reich–Sabach'13 theorem: If X is reflexive, Ψ is Legendre, (bounded, uniformly Fréchet differentiable, and totally convex) on bounded subsets of X, $\lim_{\|x\|_X\to\infty} \Psi(x)/\|x\|_X = \infty$, $\varnothing \neq K \subseteq \mathrm{int}(\mathrm{efd}(\Psi))$, $\{T_1, \ldots, T_n\}$ are LSQ(Ψ) functions $K \to K$ such that $\widehat{F} := \bigcap_{i=1}^n \widehat{\mathrm{Fix}}(T_i) \neq \varnothing$, and $T := T_1 \circ \ldots \circ T_n$, then:
  1. $\widehat{\mathrm{Fix}}(T) \subseteq \widehat{F}$,
  2. if $\widehat{\mathrm{Fix}}(T) \neq \varnothing$ then T is LSQ(Ψ).

- A function $f : X \to]-\infty, \infty]$ on a Banach space X is called totally convex at $x \in \mathrm{efd}(f)$ iff $\inf\{D_f(y, x) \mid y \in \mathrm{efd}(f),\ \|y - x\|_X \geq 0\} > 0\ \forall t \in]0, \infty[$ [Butnariu–Censor–Reich'97].

RPK'17'19: Cats of nonlinear Brègman nonexpansive operations

- As a result, we obtain the following categories of nonlinear Brègman nonexpansive operators: $\texttt{CN}(\Psi)$, $\texttt{LSQ}(\Psi)$, $\texttt{Cvx}(\Psi)$, $\texttt{Aff}(\Psi)$
- If the defining conditions of LSQ(Ψ) are assumed, and, additionally, $\mathrm{efd}(\Psi) = X$, then $\texttt{Cvx}(\Psi)$ embeds as a subcategory of $\texttt{LSQ}(\Psi)$, via $\widehat{\mathrm{Fix}}(\mathfrak{P}^{D_\Psi}_{Q_1} \cap \mathfrak{P}^{D_\Psi}_{Q_2}) = \mathrm{Fix}(\mathfrak{P}^{D_\Psi}_{Q_1} \cap \mathfrak{P}^{D_\Psi}_{Q_2}) = \mathrm{Fix}(\mathfrak{P}^{D_\Psi}_{Q_1}) \cap \mathrm{Fix}(\mathfrak{P}^{D_\Psi}_{Q_2}) = Q_1 \cap Q_2$.
- Combining this with embeddings $\ell : U \to \ell(U) \subseteq \mathrm{int}(\mathrm{efd}(\Psi))$, we obtain the categories of nonlinear postquantum operations: $\texttt{CN}(\ell, \Psi)$, $\texttt{LSQ}(\ell, \Psi)$, $\texttt{Cvx}(\ell, \Psi)$, $\texttt{Aff}(\ell, \Psi)$, with morphisms determined by

$$\widetilde{T} := \ell^{-1} \circ T \circ \ell : U \to U.$$

- $\widetilde{T}$ is an implementation of Mielnik'69'73 idea of nonlinear transmitter, although with a key difference, that we deal with ℓ-convex ℓ-closed sets.
- Following Chencov's approach, inner groupoids in the above categories are interpreted as equivalence of information models, with the corresponding notion of D_Ψ-deficiency of two Θ-parametrised models M_1 and M_2 defined as $\delta_{D_\Psi}(M_2, M_1) := \inf_{T \in Hom(M_1, \cdot)} \sup_{\theta \in \Theta} D_\Psi(\theta_2(\theta), T(\theta_1(\theta)))$.
- A composition of the embedding functor $\iota_{\ell,\Psi} : \texttt{Cvx}(\ell, \Psi) \to \texttt{LSQ}(\ell, \Psi)$ with the forgetful functor $\mathrm{Fix}_{\ell,\Psi} : \texttt{LSQ}(\ell, \Psi) \to \texttt{Cvx}(\ell, \Psi)$ (defined by attributing $\mathfrak{P}^{D_\Psi}_{\mathrm{Fix}(T)}$ to each $T \in \mathrm{Arr}(\texttt{LSQ}(\ell, \Psi))$) determines a monad $\mathrm{Fix}_{\ell,\Psi} \circ \iota_{\ell,\Psi}$ on the category $\texttt{Cvx}(\ell, \Psi)$.

Overview

- In our setting, an information state space is given by any set Z which admits a bijection ℓ into a subset of some reflexive Banach space X, such that $\ell(Z)$ is convex and closed in X.
- The objects of our categories are ℓ-closed ℓ-convex sets, which do not need to be convex (resp., normalised) in terms of the linear (resp., norm) structure of a [base] normed space.
- The good behaviour of inference (information processing) morphisms plays thus a more fundamental role than the availability of probabilistic interpretation of states.
- There is no need to restrict the domain of ℓ (and thus of D_Ψ to base norm spaces. We did it only to show the backwards compatibility and utility of our framework for the use in the postquantum ("convex operational"/"generalised probabilistic") setting.
- While the shift from commutative to noncommutative and nonassociative integration theory makes the notion of expectaton/integral more fundamental than the notion of probability/measure, the shift from linear CPTP maps and $D_\mathfrak{f}$ to nonlinear $\mathrm{LSQ}(\ell, \Psi)$ maps and D_Ψ makes the notion of an information processing (inference) more fundamental than interpretation of an information state as an integral (or as an element of a generating cone of a base norm space).
- This follows a general category-theoretic feature of proritising the objects and morphisms over the globally defined points and membership relation.
- As a result, we obtain a setting for an information theory (and, in particular, resource theories) which generally does not require spectral theory, probabilities, or integration.

Brègman nonexpansive nonlinear resource theories [RPK'19]

- Let U = a state space
(e.g., a positive generating cone or a base of a base norm space)
- A resource theory of states: a triple (P, Q, R), where
  - $P :=$ a submonoid of endomorphisms of U = free operations
  - $Q := \{\phi \in U \mid \forall\psi \in U\ \ \exists p \in P\ \ p(\psi) = \phi\}$ = set of free states if $P(Q) \subseteq Q$
  - $R := \{r : U \to \mathbb{R}^+ \mid (r \circ p)(\phi) \leq r(\phi)\ \ \forall\phi \in U\}$ = resource monotones
- Usually the free operations are assumed to be linear, and, in quantum case, CPTP. Here we provide well-defined nonlinear examples:
- Ex.1. If $\mathcal{T}$ is a monoid of $\mathrm{CN}(\ell, \Psi)$ ℓ-operations
s.t. $Q_{\mathcal{T}} := \{\phi \in U \mid \forall\psi \in U\ \ \exists t \in \mathcal{T}\ \ t(\psi) = \phi\} \neq \varnothing$ is ℓ-closed ℓ-convex
then $D_{\mathcal{T}} := \inf_{\phi \in Q_{\mathcal{T}}}\{D_\Psi(\phi, \cdot\,)\}$ is a resource monotone
and $(\mathcal{T}, Q_{\mathcal{T}}, \{D_{\mathcal{T}}\})$ is a nonlinear resource theory.
- Ex.2. If $\mathcal{T}$ is a monoid of $\mathrm{LSQ}(\ell, \Psi)$ ℓ-operations on ℓ-closed ℓ-convex $K \subseteq \mathcal{N}_\star^+$
s.t. $\bigcap_{i=1}^n \widehat{\mathrm{Fix}}(T_i) \neq \varnothing$ and $\widehat{\mathrm{Fix}}(T_1 \circ \ldots \circ T_n) \neq \varnothing\ \forall n \in \mathbb{N}\ \forall\{T_1, \ldots, T_n\} \subseteq \mathcal{T}$,
then $D_\Psi(\phi, \cdot\,)$ is a resource monotone for any $\phi \in \widehat{\mathrm{Fix}}(\mathcal{T})$
and $(\mathcal{T}, \widehat{\mathrm{Fix}}(\mathcal{T}), \bigcup_{\phi \in \widehat{\mathrm{Fix}}(\mathcal{T})}\{D_\Psi(\phi, \cdot\,)\})$ is a nonlinear resource theory.
- Ex.3. For any fixed choice of $L \in \mathrm{Ob}(\mathrm{Cvx}(\ell, \Psi))$, let $\mathcal{T}$ be given by the family of all $\mathfrak{P}^{D_\Psi}$ onto ℓ-closed ℓ-convex sets containing L. Then $(\mathcal{T}, L, \bigcup_{\phi \in L}\{D_\Psi(\phi, \cdot\,)\})$ is a nonlinear resource theory.

D_β-informations on noncommutative Banach spaces [RPK'20]

The function $\Psi_\beta(x) = \|x\|_X^{1/\beta}$, $\beta \in]0,1[$, is:

1. totally convex in any uniformly convex X [Butnariu–Iusem–Resmerita'00];
2. Legendre for any uniformly Fréchet differentiable and uniformly convex X [Bauschke–Borwein–Combettes'01].

- Hence: if X is uniformly convex and uniformly Fréchet differentiable, then $\Psi = \Psi_\beta$, satisfies conditions for composability of LSQ(Ψ).

Theorem [RPK'20]

Any noncommutative Banach space $L(\mathcal{N},\tau)$ determined by the uniformly convex symmetric function space L s.t. $L(\mathcal{N},\tau)^\star$ is determined by uniformly convex $L^\star$:

1. is naturally equipped with a family $\widetilde{D}_\beta$ of Brègman informations, determined by Ψ_β,
2. induces well defined categories $\mathtt{CN}(\Psi_\beta)$, $\mathtt{LSQ}(\Psi_\beta)$, $\mathtt{Cvx}(\Psi_\beta)$, $\mathtt{Aff}(\Psi_\beta)$,
3. for any bijective mapping $\ell : \mathcal{N}_\star^+ \to L(\mathcal{N},\tau)^+$ it induces a corresponding family of Brègman informations on $\mathcal{N}_\star$ together with the corresponding categories.

Proof: Combining the above theorems with Sukochev'86/(Dodds)$^{\otimes 2}$–de Pagter'93'14 and Krygin–Sukochev–Chilin'91 theorems.

RPK'20: quantum $D_{\Upsilon,\beta}$-informations and Jordan $D_{\gamma,\beta}$-informations

- In particular, the conditions of the above theorem hold for noncommutative Orlicz spaces $(L_\Upsilon(\mathcal{N},\tau),\|\cdot\|_\Upsilon)$, where $\mathcal{N}$ has type II_∞, Υ and Υ^{Y} are uniformly convex Orlicz functions satisfying $\triangle_2$ condition. (For other semi-finite types of $\mathcal{N}$ there are corresponding, slightly different, conditions.)
- By introducing noncommutative Kaczmarz map $\ell_\Upsilon : \mathcal{N}_\star \ni \phi = u_\phi|\phi| \mapsto u_\phi \Upsilon^{-1}(|\phi|) = u_\phi \Upsilon^{-1}(\Delta_{\phi,\tau}) \in L_\Upsilon(\mathcal{N},\tau)$, we obtain a family of $D_{\beta,\Upsilon}$ informations (and corresponding categories) on preduals of semi-finite W*-algebras. (An extension to all predual $\mathcal{N}_\star$ is due to uniqueness of polar decomposition, combined with replacing $[\![\cdot,\cdot]\!]_{X\times X^\star}$ with $\mathrm{re}\,[\![\cdot,\cdot]\!]_{X\times X^\star}$ in the definition of $\widetilde{D}_\Psi$.)
- Under restriction to $\Upsilon(x) = x^{1/\gamma}$, $\gamma \in]0,1[$, corresponding to noncommutative $L_{1/\gamma}$ spaces, the condition of semi-finiteness of W*-algebras is obsolete, due to uniform convexity of any $L_{1/\gamma}(\mathcal{N})$ [Terp'81, Masuda'83, Kosaki'84]. The corresponding family of $D_{\beta,\gamma}$-informations (as well as the corresponding categories) is well-defined on preduals of arbitrary W*-algebras.
- By combining B–I–R'00 and B–B–C'01 theorems with uniform Fréchet differentiability and uniform convexity of $L_{1/\gamma}(A,\tau)$ spaces over semi-finite JBW-algebras A [Iochum'84'86, Ayupov'86], and introducing the nonassociative Mazur map $\ell_{1/\gamma} : L_1(A,\tau) \ni |\phi| \circ s_\phi \mapsto |\phi|^\gamma \circ s_\phi \in L_{1/\gamma}(A,\tau)$, $s_\phi^2 = \mathbb{I}$, we obtain a family of $D_{\beta,\gamma}$ informations (and corresponding categories) on preduals of semi-finite JBW-algebras.

Topics omitted in this talk

- Right Brègman projections, right Brègman nonexpansive operators, etc., together with categorical equivalence of right and left categories [RPK'20].
- Smooth information geometric side of the theory:
  - $\widetilde{D}_\Psi$ over sets C which are dually affine (i.e., affine in X and $\mathfrak{D}^{\mathrm{G}}\Psi$-affine in $X^\star$), gives naturally rise to doubly (flat, torsion-free, autoparallel) affine geometry $(\mathcal{M}, \nabla, \nabla^\dagger)$ over $\ell^{-1}(C)$ [RPK'20],
  - D_Ψ-projections onto dually afine sets coincide with the geodesic projections,
  - under additional assumption on D_Ψ (very strict convexity, i.e. positive definiteness of hessian of Ψ), third order Taylor expansion of D_Ψ gives rise to dually flat dually torsion-free Norden–Sen geometry, known as hessian geometry, which is a special case of the above geometry,
  - these geometries allow to describe the Jaynes–Mitchell approach to source renormalisation [Favretti'07],
  - topos-theoretic algebraisation (and representation, using categories of presheaves of the Postnikov–Sikorski spaces) of these geometries [RPK'19].
- Epistemic adjointness (categorical (co)monadic resource theory) [RPK'12'16'19]:
  - Two categories: experimental design $\mathtt{ExpDes}$, inductive inferences/information processings $\mathtt{IndInf}$. Model construction as semantics functor $\mathtt{ExpDes} \to \mathtt{IndInf}$, predictive verification as syntax functor $\mathtt{IndInf}$. "Epistemic" comonad E on $\mathtt{IndInf} \to \mathtt{ExpDes}$ implementing abstractly the above relationship. Monad J on $\mathtt{IndInf}$ implementing free operations. $(\mathtt{IndInf}, E, J)$ as a categorical resource theory
  - Epistemic comonad $E_{\ell,\Psi}$ on $\mathtt{Cvx}(\ell, \Psi)$ induced by a ℓ-convex-closure functor on subsets of underlying set, combined with the forgetful functor.
  - A triple $(\mathtt{Cvx}(\ell, \Psi), E_{\ell,\Psi}, \mathrm{Fix}_{\ell,\Psi} \circ \iota_{\ell,\Psi})$ as an example of categorical resource theory
- Postjaynesian interpretation of all of this framework [RPK'10+...].

References

- RPK, 2017/19/20, Brègman relative entropies and nonlinear information theory, arXiv:1710.01837 (v1: 2017, v2: 2019, v3: 2020 [forthcoming]).
- RPK, 2016, Local quantum information dynamics, arXiv:1605.02063.
- M.I.Munk-Nielsen, 2015, Quantum measurements from entropic projections, M.Sc. Thesis, Perimeter Institute, Waterloo, Canada.
- RPK, 2014, Lüders' and quantum Jeffrey's rules as entropic projections, arXiv:1408.3502.
- F. Hellmann, W. Kamiński, RPK, 2014, Quantum collapse rules from maximum relative entropy principle, New J. Phys. **18** (2016), 013022. arXiv:1407.7766.
- RPK, 2013, W^*-algebras and noncommutative integration, arXiv:1307.4818.
- RPK, 2013, Information geometric foundations of quantum theory, Ph.D. thesis, Institute for Theoretical Physics, University of Warsaw.
- RPK, 2011, The general form of γ-family of quantum relative entropies, Open Systems and Information Dynamics **18**, 191–221. arXiv:1106.2225.
- RPK, 2010, Quantum theory as inductive inference, in: Mohammad-Djafari A., Bercher J., Bessière P. (eds.), *Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, AIP Conf. Proc. **1305**, 28–35. arXiv:1009.2423

本日、
お時間を割いて頂きありがとうございます。